N° 92F0138MIF au catalogue — N° 001
ISSN : 1481-1758
ISBN : 978-0-662-73088-0

Document de recherche

Série de documents de travail de la géographie

Mise en correspondance des codes postaux et des régions géographiques

par Kaveri Mechanda and Henry Puderer

Division de la géographie
Édifice Jean-Talon, 3e étage, Ottawa, K1A 0T6

Téléphone: 613-951-0437 Télécopieur: 613-951-0569

Statistique Canada Statistics Canada

Canada

Série de documents de travail de la géographie

Mise en correspondance des codes postaux et des régions géographiques

Par

Kaveri Mechanda et Henry Puderer

92F0138MIF — N° 001
ISSN : 1481-1758
ISBN : 978-0-662-73088-0

Division de la géographie, Statistique Canada
Édifice Jean-Talon, 3e étage, Ottawa, K1A 0T6

Comment obtenir d'autres renseignements :
Téléphone : 613-951-0437
Télécopieur : 613-951-0569
Renseignements par courriel : geohelp@statcan.ca

Février 2007

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'industrie, 2007

Tous droits réservés. Le contenu de la présente publication électronique peut être reproduit en tout ou en partie, et par quelque moyen que ce soit, sans autre permission de Statistique Canada, sous réserve que la reproduction soit effectuée uniquement à des fins d'étude privée, de recherche, de critique, de compte rendu ou en vue d'en préparer un résumé destiné aux journaux et/ou à des fins non commerciales. Statistique Canada doit être cité comme suit : Source (ou « Adapté de », s'il y a lieu) : Statistique Canada, année de publication, nom du produit, numéro au catalogue, volume et numéro, période de référence et page(s). Autrement, il est interdit de reproduire le contenu de la présente publication, ou de l'emmagasiner dans un système d'extraction, ou de le transmettre sous quelque forme ou par quelque moyen que ce soit, reproduction électronique, mécanique, photographique, pour quelque fin que ce soit, sans l'autorisation écrite préalable des Services d'octroi de licences, Division des services à la clientèle, Statistique Canada, Ottawa, Ontario, Canada K1A 0T6.

Remerciements

Les auteurs aimeraient remercier Belinda Ha pour son analyse et ses tabulations en SAS, Robert Parenteau, Sandra Porter et Joseph Kresovic pour la réduction de contenu et également Cathy Caldwell pour l'implantation de géocodage automatisé.

Série de documents de travail de la géographie

La Série de documents de travail de la géographie vise à stimuler des discussions sur une variété de sujets reliés au travail conceptuel, méthodologique et technique à l'appui de l'élaboration et de la diffusion des données, des produits et des services de la division. Nous encourageons les lecteurs de la série à communiquer avec la Division de la géographie pour lui fournir leurs commentaires et suggestions.

This publication is available in English upon request (Catalogue no. 92F0138MIE).

Table des matières

Aperçu **4**

Résumé **5**

Introduction **6**

1.0 Typologie des codes postaux et localisation **8**

1.1 Le code postal et la localisation 8
1.2 Données sur les adresses et la livraison tirées des fichiers de la SCP 10
1.3 Fichier Recherche d'adresses 10
1.4 Fichiers de localisation des installations de livraison 18
1.5 Type d'enregistrement et type de mode de livraison à partir du Fichier sur les modes de livraison 19
1.6 Considérations relatives au géocodage fondé sur les données de la SCP 20

2.0 Utilisateurs des codes postaux et géographie **21**

2.1 Personnes et codes postaux déclarés dans le recensement 22
2.2 Identification du secteur de recherche pour le géocodage 24
2.3 Géocodage des codes postaux directement à partir des données du recensement 24
2.4 Recommandations pour le géocodage fondé sur les réponses du recensement 27

3.0 Règles de géocoage et mesures de la qualité **28**

3.1 Secteurs de service et secteurs de recherche 28
3.2 Géocodage des codes postaux selon le CPType 29
3.3 Mesures de la qualité 31

Conclusion **37**

Aperçu

Statistique Canada crée des fichiers qui fournissent le couplage entre les codes postaux et les régions géographiques selon lesquelles les données statistiques sont diffusées. En couplant les codes postaux et ses régions géographiques, Statistique Canada facilite l'extraction et l'agrégation subséquente de données pour certaines régions géographiques de fichiers mis à la disposition des utilisateurs. Les utilisateurs peuvent alors totaliser les données de Statistique Canada pour leurs régions et d'autres données pour ces mêmes régions de manière à établir un profil statistique combiné de ces régions.

La méthodologie utilisée par Statistique Canada pour le couplage des codes postaux et des régions géographiques a été questionnée. Pour régler la situation, Statistique Canada a décidé de créer un cadre conceptuel pour l'élaboration des règles régissant le couplage des codes postaux et des régions géographiques de Statistique Canada. Le présent document de travail comprend le cadre conceptuel et a pour objectif ce qui suit :

- étudier les concepts et les descripteurs de codes postaux utilisés par la Société canadienne des postes, du point de vue de leur typologie et de leur occurrence;
- examiner comment les codes postaux sont utilisés et élaborer des secteurs de service de code postal fondés sur cette utilisation;
- établir les règles opérationnelles régissant le géocodage ou le couplage des codes postaux et des régions géographiques normalisées, et fournir des indicateurs de la qualité des données et d'autres métadonnées.

La méthodologie qui y est décrite servira de fondement pour le couplage des codes postaux aux régions géographiques du Recensement de 2006. Les utilisateurs des produits de Statistique Canada qui font appel aux codes postaux sont invités à formuler des commentaires sur le présent document.

Le lien entre les codes postaux et les régions géographiques de Statistique Canada est actuellement assuré grâce au Fichier de conversion des codes postaux (FCCP) et au Fichier des codes postaux par circonscription électorale fédérale (FCPCEF). On s'attend à ce que les règles de géocodage et les recommandations relatives aux indicateurs de qualité puissent être utilisées pour créer ces produits. Ces nouvelles règles ont été mises à l'essai dans le cadre du couplage des régions géographiques du Recensement de 2001 et ont entraîné la révision de plus de 40 000 enregistrements ou environ 2 % des enregistrements du FCCP. Les essais effectués jusqu'à maintenant montrent qu'environ 90 % des codes postaux peuvent être géocodés grâce aux nouvelles règles de géocodage établies dans le cadre du présent document de travail. Des recherches plus poussées sont nécessaires à l'égard du géocodage des 10 % de codes postaux qui restent; on s'attend à ce que les données du Recensement de 2006 aident au géocodage de certains d'entre eux. On prévoit aussi que les codes postaux du FCCP et du FCPCEF seront géocodés en fonction des régions géographiques du Recensement de 2006, selon les nouvelles règles de géocodage.

Résumé

Statistique Canada crée des fichiers qui fournissent le couplage entre les codes postaux et les régions géographiques selon lesquelles les données statistiques sont diffusées. En couplant les codes postaux et ses régions géographiques, Statistique Canada facilite l'extraction et l'agrégation subséquente de données pour certaines régions géographiques de fichiers mis à la disposition des utilisateurs. Les utilisateurs peuvent alors totaliser les données de Statistique Canada pour leurs régions et d'autres données pour ces mêmes régions de manière à établir un profil statistique combiné de ces régions.

La méthodologie utilisée par Statistique Canada pour le couplage des codes postaux et des régions géographiques a été questionnée. Pour régler la situation, Statistique Canada a décidé de créer un cadre conceptuel pour l'élaboration des règles régissant le couplage des codes postaux et des régions géographiques de Statistique Canada. Le présent document de travail expose le cadre conceptuel et les règles de géocodage. La méthodologie qui y est décrite servira de fondement pour le couplage des codes postaux aux régions géographiques du Recensement de 2006. Les utilisateurs des produits de Statistique Canada qui font appel aux codes postaux sont invités à formuler des commentaires sur le présent document.

Introduction

Le présent document comporte deux objectifs : 1) établir un cadre conceptuel définissant le rapport entre les codes postaux et les régions géographiques normalisées; et 2) définir par la suite des règles pour le couplage des codes postaux et des *régions géographiques normalisées*.[1]

Les codes postaux sont gérés par la Société canadienne des postes (SCP) et servent au tri et à la livraison efficace du courrier. Ils ne sont pas créés comme unités pour l'analyse ou la mise en correspondance de caractéristiques de populations, d'entreprises ou de logements. Les produits fondés sur les codes postaux créés par Statistique Canada servent de lien entre les codes postaux et les régions géographiques normalisées, afin de permettre diverses utilisations des données. De façon générale, les utilisations des produits fondés sur les codes postaux peuvent être caractérisées de la façon suivante :

1. association des codes postaux et des régions géographiques normalisées (pour l'extraction de données, la mise en correspondance, l'établissement de profils et/ou l'analyse);
2. la sélection d'***unités statistique*** selon les *régions géographiques normalisées* pour l'échantillonnage, les études de marché, etc.

Unité statistique

Une unité statistique est l'unité d'observation ou de mesure pour laquelle des données sont recueillies ou calculées. L'unité statistique de base pour les statistiques sociales est la personne, qui sert d'unité d'analyse dans le cadre de la plupart des programmes de statistiques sociales. Parmi les autres unités utilisées dans les statistiques sociales figurent le *logement* et le *ménage*.

L'association entre les codes postaux et les régions géographiques normalisées est actuellement comprise dans les produits suivants : 1) le Fichier de conversion des codes postaux (FCCP) et 2) le Fichier des codes postaux par circonscriptions électorales fédérales (FCPCÉF). Le FCCP permet d'assurer le couplage des codes postaux et de la plupart des régions géographiques normalisées, y compris les entités détaillées suivantes : *côté d'îlot*, *îlot de diffusion* et *aire de diffusion*. Le FCCP est le principal produit répondant aux exigences des applications mentionnées précédemment.

Au fil des ans, un certain nombre de problèmes se sont posés relativement à la qualité du FCCP, ce qui a entraîné la remise en question des concepts fondamentaux qui sous-tendent la façon dont les codes postaux sont couplés aux régions géographiques normalisées. Il s'agit de problèmes comme l'utilisation appropriée des données pour le couplage des codes postaux, la précision et l'exactitude des couplages, et l'attribution d'*indicateurs de lien unique*. Le suivi de ces problèmes de qualité du FCCP a fait augmenter les coûts déjà élevés du géocodage des codes postaux. Ces problèmes ont fait ressortir la nécessité d'examiner et d'établir un cadre conceptuel définissant le rapport entre les codes postaux et les régions géographiques normalisées, en vue de rétablir les règles opérationnelles servant au couplage des codes postaux et des régions géographiques normalisées, ainsi qu'à l'élaboration subséquente des produits.

Le couplage des codes postaux et des régions géographiques a posé des difficultés dans les cas suivants :

- Il n'existait pas de rapport explicite entre le nom de municipalité/collectivité de code postal dans les fichiers de la SCP et les régions géographiques normalisées. Ce rapport a à l'occasion fait l'objet d'erreurs d'interprétation qui ont entraîné des couplages incorrects.
- Étant donné que les empreintes géographiques types des différents genres de codes postaux n'étaient pas bien comprises, les codes postaux ont parfois été couplés à un trop

1. Les termes en *italiques* sont définis dans le glossaire

grand nombre (ou un trop petit nombre) de régions géographiques, ce qui a entraîné un surdénombrement (ou un sous-dénombrement).

- Dans certains cas, le surdénombrement mentionné précédemment a fait en sorte qu'un code postal a été couplé à un trop grand nombre de petites régions géographiques (comme des îlots), ce qui implique que le niveau de précision était inapproprié.

Le présent document de travail repose sur l'hypothèse selon laquelle une meilleure utilisation des données sur les codes postaux recueillies dans le cadre du recensement ainsi que des données des fichiers de la SCP contribuera à résoudre les problèmes décrits précédemment et permettra d'obtenir des couplages plus précis entre les codes postaux et les régions géographiques normalisées.

Des règles doivent être élaborées pour géocoder les codes postaux selon les régions géographiques normalisées de Statistique Canada, afin d'améliorer le couplage. Le géocodage peut être défini comme le processus d'attribution d'identificateurs géographiques (codes), en vue de la mise en correspondance des traits physiques et des enregistrements de données. Les codes géographiques en résultant permettent le couplage géographique des données. Par exemple, l'une des façons de géocoder une *adresse municipale* au Canada consiste à localiser d'abord la région générale, puis l'immeuble particulier comportant cette adresse municipale. Étant donné que les mêmes noms de rue ou adresses municipales peuvent se retrouver dans de nombreuses régions au Canada, il est essentiel de pouvoir limiter la recherche d'une combinaison de nom de rue et d'adresse municipale à un lieu particulier au Canada. Le terme « *secteur de recherche* » sera utilisé dans le présent document pour désigner la région géographique à l'intérieur de laquelle le géocodage peut être effectué à partir des noms de rue et des adresses municipales. Le processus de géocodage d'une adresse municipale peut par conséquent être décrit de la façon suivante :

Établir le *secteur de recherche*

↓

Trouver le secteur de recherche dans le Fichier du réseau routier

↓

Trouver la rue dans le secteur de recherche du Fichier du réseau routier

↓

Trouver l'adresse dans une tranche d'adresses d'une rue

↓

Établir les codes géographiques

(Les codes géographiques utilisés sont le côté d'îlot, l'îlot de diffusion et l'aire de diffusion)

Le processus de géocodage décrit précédemment servira de modèle général pour établir les règles opérationnelles de géocodage des codes postaux. Le processus de géocodage des codes postaux permet ultimement de créer le lien entre les codes postaux et les codes géographiques, soit le côté d'îlot, l'îlot de diffusion et l'aire de diffusion. Le lien entre les codes postaux et toutes les autres régions géographiques normalisées est fondé sur ce géocodage initial.

Le présent document de travail vise les objectifs suivants :

1. Étudier les concepts et les descripteurs de codes postaux utilisés par la Société canadienne des postes, du point de vue de leur typologie et de leur occurrence, à titre de première étape en vue de comprendre la façon dont les codes postaux sont mis à la disposition des utilisateurs. Cette étude fait partie de la section 1.
2. Examiner comment les codes postaux sont utilisés et élaborer des *secteurs de service de code postal* fondés sur les modèles d'utilisation, afin que ces secteurs de service puissent servir comme secteurs de recherche pour le géocodage. Cette étape est comprise dans la section 2.
3. Établir les règles opérationnelles régissant le géocodage ou le couplage des codes postaux et des régions géographiques normalisées, et fournir des indicateurs de la qualité des données et d'autres métadonnées. Cette étape est comprise dans la section 3.

On s'attend à ce que les règles de géocodage et les recommandations d'indicateurs de la qualité puissent être utilisées pour créer les produits.

1.0 Typologie des codes postaux et localisation

L'examen de la structure et de la typologie des codes postaux constitue la première étape en vue de comprendre comment ils peuvent être liés à des régions géographiques normalisées. Certaines des définitions pertinentes, de même que les types et les occurrences, sont examinées dans la présente section. Cela sert de base aux règles opérationnelles de géocodage des codes postaux (voir la section 3.0).

1.1 Le code postal et la localisation

Les codes postaux sont gérés par la SCP. Ils ont été conçus pour aider à trier rapidement le courrier et à assurer la livraison plus efficace des lettres, colis et autres envois par la SCP. Un code postal peut être lié à différents types de points de livraison, y compris des boîtes aux lettres résidentielles, des boîtes postales multiples, ainsi que des cases postales de bureaux de poste (voir www.postescanada.ca).

Les caractères qui constituent le code postal sont généralement représentatifs de l'utilisation prévue (livraison) du code postal. La Société canadienne des postes définit un code postal de la façon suivante :

> Une combinaison alphanumérique de six caractères (ANA NAN) attribuée à une ou plusieurs adresses postales. Le code postal fait partie intégrante de toutes les adresses postales au Canada et est nécessaire pour le traitement mécanisé du courrier. Les codes postaux servent aussi à identifier les diverses installations de traitement et de livraison de la SCP (Société canadienne des postes, 2005a, page 30).

Les trois premiers caractères du code postal (combinaison alphanumérique de « ANA ») correspondent à la *région de tri d'acheminement (RTA)*. Cette dernière est définie de la façon suivante :

> La région de tri d'acheminement ou RTA représente une zone précise au sein d'une province ou d'une zone géographique importante. La région de tri d'acheminement constitue le fondement du tri primaire du courrier d'acheminement (Société canadienne des postes, 2007, section B).

Les trois derniers caractères du code sont appelés unité de distribution locale (UDL). Ils permettent la création de codes postaux à l'intérieur du secteur de service de la RTA.

Chacune des premières lettres du code postal est réservée à une province / un territoire particulier du pays (voir l'annexe A pour plus de détails). De façon plus particulière, selon les dossiers de la SCP, le premier caractère correspond à la province/au territoire de l'*installation de livraison* à partir de laquelle le service lié au code postal est assuré. Le deuxième caractère du code postal correspond à la couverture de ce code. Un code postal comportant un « 0 » comme deuxième caractère est classé comme « rural », et tous les autres codes postaux sont considérés comme urbains par la Société canadienne des postes.[2] Toutefois, cela ne signifie pas que le code postal correspond aux « régions rurales » ou « régions urbaines » définies selon les chiffres de population du recensement courant.[3,4] Ils peuvent être désignés comme des codes postaux desservant des secteurs de *livraison rurale* et des secteurs de *livraison urbaine*, afin d'éviter la confusion avec les termes « régions rurales » et « régions urbaines » utilisés par Statistique Canada. Par exemple, le Nouveau-Brunswick ne compte aucun secteur de livraison rurale, même si Statistique Canada considère cette province comme rurale pour une large part. Dans le cas des codes postaux liés à des secteurs de livraison rurale, le Guide des postes du Canada indique ce qui suit : « ... lorsqu'ils sont combinés à la région de tri d'acheminement, les trois derniers caractères désignent une collectivité rurale précise » (Société canadienne des postes, 2007, section B, chapitre 3, page 10). Les collectivités doivent être localisées sur une carte, afin de pouvoir être géocodées selon les régions géographiques normalisées.

Des cartes sont disponibles auprès de la SCP et correspondent au secteur représenté par une RTA. Toutefois, les personnes vivant dans une province peuvent utiliser un code postal lié à une autre province, et les personnes vivant dans une RTA peuvent utiliser le code postal d'une autre RTA. Cela signifie que le secteur desservi par le code postal ou la RTA peut être différent des limites administratives créées par la Société canadienne des postes. En outre, l'échelle de ces cartes (1:50 000 à 1:12 000 000) rend difficile le lien entre les limites des RTA et le réseau routier plus détaillé du cadre géographique de Statistique Canada. Même si ces cartes sont utiles pour la localisation de la RTA à l'intérieur d'une province, elles ne nous renseignent pas sur la portée de la RTA en rapport avec un réseau routier disponible dans le public, et ne comprennent pas non plus les renseignements plus détaillés nécessaires pour le géocodage des codes postaux. Toutefois, les cartes des itinéraires de facteurs (IF) disponibles récemment sur Internet (généralement pour les secteurs de livraison urbaine) présentent les limites des RTA dans le contexte du réseau routier. Elles peuvent être plus utiles que les cartes des RTA pour déterminer les limites de nombreuses RTA. Toutefois, elles ne permettent pas de délimiter le secteur desservi par un code postal.

Le code postal proprement dit n'est pas suffisant pour déterminer l'emplacement exact et la couverture du secteur de service. La RTA peut être utilisée pour localiser le secteur général desservi par les codes postaux pertinents. Des renseignements provenant d'autres sources sont nécessaires pour associer le code postal et des régions géographiques normalisées particulières.

---

2. La classification de la SCP est considérée comme une composante de la préparation du courrier. La classification est fondée sur le deuxième caractère du code postal, de la façon suivante :
   - code postal urbain : chiffres de 1 à 9 (E2J). Les codes postaux urbains sont généralement desservis par un facteur ou par des boîtes postales communautaires; ou
   - code postal rural : chiffre de 0 (zéro) (A0A). Les codes postaux ruraux sont desservis par des entrepreneurs des routes rurales et/ou des comptoirs postaux.

3. Dans le présent document de travail, le terme « recensement » désigne le recensement de la population.
4. Statistique Canada définit les régions urbaines et rurales de la façon suivante :
   Une région urbaine compte au minimum 1 000 habitants et à une densité de population d'au moins 400 personnes par kilomètre carré, selon les chiffres du recensement courant de la population. L'ensemble du territoire situé à l'extérieur des régions urbaines est classé comme rural. Ensemble, les régions urbaines et rurales couvrent tout le territoire canadien.

1.2 Données sur les adresses et la livraison tirées des fichiers de la SCP

La SCP fournit des renseignements sur les codes postaux dans le contexte des différents aspects de la livraison du courrier. Ces renseignements sont pertinents du point de vue de la façon dont ces codes postaux se rejoignent géographiquement et dont ils peuvent être consultés par les personnes et les institutions qui utilisent le système postal. Les typologies et les données de localisation décrites dans le présent document sont tirées des données fournies sur une base mensuelle par la SCP à Statistique Canada, dans les fichiers suivants.

1. Fichier Recherche d'adresses

   Ce fichier comprend tous les codes postaux actifs et toutes les adresses postales valides au pays (Société canadienne des postes, 2005a). Une version mise à jour de ce fichier est disponible sur une base mensuelle auprès de la SCP, et elle comprend tous les codes postaux qui sont considérés comme « actifs » pour le service postal à ce moment-là. Environ 97 % des codes postaux du fichier sont liés à la livraison à une adresse municipale[5]. Ces codes postaux peuvent être associés à plus d'un enregistrement de tranche d'adresses dans le fichier. Ces enregistrements comprennent un nom de collectivité utilisé par la SCP, un nom de rue et une tranche d'adresses municipales. (Les différents types d'enregistrements compris dans le Fichier Recherche d'adresses et le nom de *collectivité de la SCP* sont décrits dans la section 1.3.)
2. Fichier sur les chefs de ménage

   Ce fichier est conçu pour fournir des renseignements sur le nombre de logements et d'entreprises desservis par les installations de livraison. Il comprend toutes les installations de livraison rurales et urbaines au Canada (Société canadienne des postes, 2005b). Il comprend en outre les adresses de certaines de ses installations. Ces adresses peuvent être utiles pour le géocodage des codes postaux en fonction des installations dont ils relèvent. (L'utilisation du Fichier sur les chefs de ménage est abordée à la section 1.4.)
3. Fichier sur les modes de livraison

   Ce fichier comprend les codes postaux et les modes de livraison s'y rapportant (Société canadienne des postes, 2005c).

   Il indique si la livraison est faite à une tranche d'adresses municipales, à un immeuble commercial ou à une case postale de bureau de poste. (Les types de mode de livraison sont abordés de façon plus détaillée à la section 1.5.)

Il s'agit des trois fichiers les plus pertinents pour le géocodage des codes postaux. Parmi eux, le Fichier Recherche d'adresses, qui comprend les données sur les adresses liées à la grande majorité des codes postaux, est le plus utile pour le géocodage.

1.3 Fichier Recherche d'adresses

Le Fichier Recherche d'adresses fournit un tableau complet des codes postaux actifs (y compris les mises à jour et les 500 à 2 500 nouveaux codes postaux adoptés ou « créés » chaque mois). Parmi les 1 144 108 enregistrements figurant dans le Fichier Recherche d'adresses de juin 2006, 1 133 814 étaient « actifs » ou valides pour le service postal. Les autres codes postaux du fichier sont ceux qui sont « retirés » au cours du mois en question. Statistique Canada utilise le terme code postal retiré pour décrire tous les codes postaux qui étaient actifs à un moment donné, mais ne le sont plus. Même si la Société canadienne des postes considère uniquement les codes postaux actifs comme valides pour le service postal, les codes postaux retirés peuvent être utilisés un certain temps après leur retrait.

5. Les autres 3 % de codes postaux ne sont pas liés à des adresses municipales, mais à des cases postales de bureaux de poste et à des itinéraires de service postal.

Nom de collectivité de la SCP dans le Fichier Recherche d'adresses

L'identification de la municipalité desservie par un code postal constitue une étape initiale importante pour déterminer avec précision où le code postal devrait être géocodé. Le premier caractère du code postal (qui indique la province ou le territoire), combiné au nom de collectivité de la SCP permet parfois, mais pas toujours, de repérer la municipalité ou la *subdivision de recensement* (SDR) à partir de laquelle le code postal peut être géocodé. Lorsque la SDR peut être localisée, les données sur l'adresse municipale peuvent servir au géocodage en fonction d'une tranche d'adresses municipales.

Le terme « municipalité » dans le Fichier Recherche d'adresses ne correspond pas nécessairement à la municipalité définie par la loi au moment où la Société canadienne des postes publie les données. Il s'agit de noms utilisés au niveau local, afin de prévenir la confusion en ce qui a trait au secteur desservi. Le secteur desservi ne coïncide pas nécessairement avec les municipalités légalement constituées ayant le même non ou un nom similaire. Par exemple, lorsque des municipalités sont fusionnées, la Société canadienne des postes préserve les anciens noms (avant la fusion), jusqu'à ce qu'elle puisse établir que les différentes rues comportent un nom unique dans la nouvelle municipalité fusionnée. La RTA est aussi définie en rapport avec une municipalité de la façon suivante :

> La RTA comprend les trois premiers caractères d'un code postal. La RTA qualifie la zone, à l'intérieur d'une municipalité, qui est autorisé à utiliser le nouveau nom valide de municipalité (Société canadienne des postes, 2005a, page 33).

Ces noms de municipalités sont souvent des noms de lieux qui n'existent plus ou des noms couramment utilisés en rapport avec ces localités. Afin d'éviter toute confusion, on utilisera ci-après le terme « collectivité de la SCP ».

La comparaison des noms de municipalités utilisés en janvier 2001 illustre ce point. Les SDR de Statistique Canada comprennent toutes les municipalités définies par la loi qui existaient en date du 1er janvier de l'année du recensement (2001). Une comparaison des noms de collectivité *de la SCP* existants montre que moins de 40 % d'entre eux correspondent aux SDR. Le tableau qui suit illustre les appariements de noms de collectivités de la SCP et de noms de SDR de 2001[6] :

6. Ce taux d'appariement est fondé sur l'appariement de noms uniques de SDR à l'intérieur d'une province/territoire et des noms de collectivités de la SCP correspondant aux codes postaux d'une province/territoire (selon la première lettre du code postal). L'appariement a été effectué après uniformisation des noms en lettres majuscules, sans accent ni caractère spécial. Les chiffres du tableau constituent en fait une légère surestimation du taux d'appariement, étant donné que le type de SDR n'a pu être pris en compte dans l'appariement et ne figure pas dans le dénominateur servant au calcul du pourcentage. (Les SDR comportant le même nom, mais un type différent (ville, réserve, district municipal, etc.), comme la ville de Yarmouth et le district municipal de Yarmouth, les deux se trouvant en Nouvelle-Écosse, sont des SDR entièrement différentes. Voir le glossaire pour une définition du type de SDR.)

Tableau 1 Noms de *collectivités de la SCP* et noms de SDR

Fichier Recherche d'adresses		Appariement avec les SDR du Recensement de 2001	
Date de référence*	Nombre de *collectivités de la SCP*		
		nombre	%
Novembre 2000	6 983	2 689	38,5
Juin 2001	7 012	2 697	38,5
Décembre 2003	6 987	2 708	38,8
Décembre 2004	6 970	2 703	38,8
Juin 2005	6 961	2 701	38,8
Octobre 2005	6 958	2 707	38,9

* Les dates de référence rendent compte des codes postaux adoptés/gardés actifs pour le mois et l'année en question.
Sources de données : Société canadienne des postes, Novembre 2000 - Octobre 2005, Fichier Recherche d'adresses et Statistique Canada, Recensement de la population de 2001.
Source du tableau : Statistique Canada, 2007, Mise en correspondance des codes postaux et des régions géographiques, numéro 92F0138MWF2007001 au catalogue.

Le taux d'appariement dans le tableau qui précède est d'environ 40 %, en dépit du traitement (uniformisation) assuré pour l'augmenter. L'examen du Fichier Recherche d'adresses de juin 2001 montre les changements qui se sont produits par suite des fusions. Par exemple, Nepean et Ottawa ont été conservés comme noms de collectivités de la SCP, même si Nepean n'existe plus légalement comme municipalité par suite de la fusion. Les noms de rues et les noms de collectivités de la SCP semblent être mis à jour dans le Fichier Recherche d'adresses pour rendre compte des changements qui se sont produits au niveau municipal, une fois que la SCP a résolu la question du caractère unique de l'adresse. Cela peut expliquer l'augmentation légère des taux d'appariement sur quatre ans. Toutefois, aux fins du géocodage, un fichier de correspondance entre les noms de SDR et de collectivités de la SCP est nécessaire, afin que le géocodage puisse être ramené au niveau de la SDR en premier lieu, puis au niveau de l'adresse municipale à l'intérieur de celle-ci[7].

Malheureusement, même lorsque les noms de SDR correspondent, il faut plus de renseignements pour définir le secteur à l'intérieur duquel les données sur les adresses municipales peuvent être appariées de façon unique. Il se peut que plus d'une rue ait le même nom à l'intérieur d'une SDR. Le même code postal peut parfois être attribué à plusieurs SDR. Le géocodage possible à l'intérieur de collectivités de la SCP multiples devrait être envisagé.

Le nom de collectivité de la SCP peut être un outil utile pour localiser les municipalités où se trouve l'adresse municipale. Parfois, une correspondance peut être établie entre la SDR et le nom de collectivité de la SCP sur la base de la province/territoire et des similitudes entre les noms. Toutefois, en raison des différentes dates de référence et des objectifs visés par le nom de SDR et le nom de collectivité de la SCP, cette correspondance n'est pas toujours possible et, même lorsqu'elle est établie, ne peut être effectuée avec un niveau de confiance absolu. Des données supplémentaires sont nécessaires pour établir le secteur à l'intérieur duquel l'adresse municipale peut se retrouver.

7. Il est essentiel de trouver le secteur de service propre à un code postal pour établir un processus de géocodage automatisé. Toutefois, dans le cadre d'un processus de géocodage manuel, il existe des options de consultation de diverses cartes, et même des municipalités.

Types de codes postaux dans le Fichier Recherche d'adresses

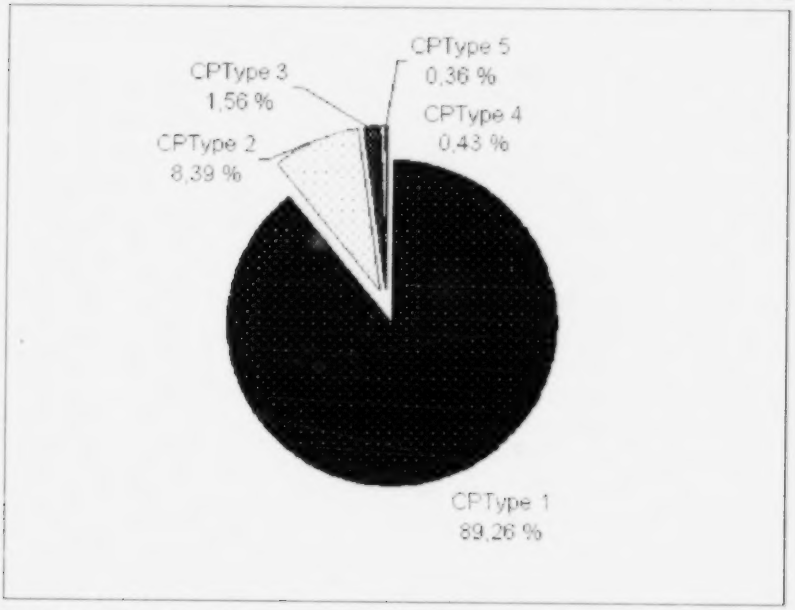
Une fois que la municipalité ou le secteur général desservi par le code postal est établi, le secteur de service propre au code postal doit être défini avant le géocodage. Les types d'enregistrements du Fichier Recherche d'adresses sont essentiels pour établir le type de *secteur de service du code postal*. La Société canadienne des postes désigne cinq types d'enregistrements dans le Fichier Recherche d'adresses[8]. Ils sont liés à cinq types différents de livraison. Ces enregistrements sont catégorisés et appelés « CPType » dans le FCCP. La typologie est la suivante[9] :

- CPType 1, Adresse municipale
  Ce type d'enregistrement contient les codes postaux des adresses municipales.
- CPType 2, Rue desservie par un service associé à un itinéraire ou *poste restante*
  Ce type d'enregistrement contient les codes postaux attribués aux adresses municipales desservies par un service associé à un itinéraire ou à la poste restante à partir d'installations de livraison – souvent en périphérie des régions urbaines.
- CPType 3, Adresse de *case postale*
  Ce type d'enregistrement contient le code postal d'une *case postale* ou d'une tranche de numéros de *cases postales* à l'intérieur d'une installation de livraison.
- CPType 4, Service associé à un itinéraire
  Ce type d'enregistrement contient les codes postaux attribués aux adresses se trouvant sur un itinéraire de *route rurale*, sur un *itinéraire motorisé* ou sur un itinéraire de service suburbain desservi à partir d'une installation de livraison.
- CPType 5, Poste restante
  Ce type d'enregistrement contient les codes postaux attribués à la poste restante à une installation de livraison.

Un code postal peut être lié à plus d'un CPType d'enregistrement. Chaque enregistrement du Fichier Recherche d'adresses de juin 2006 comporte un code postal et un type particulier. Comme le montre le graphique ci-dessous, la grande majorité des enregistrements sont de CPType 1 ou 2.

8. Ces renseignements sont fondés sur une analyse des données sur les codes postaux tirées de divers documents de la Société canadienne des postes dans le site Internet de la SCP : http://www.postescanada.ca/ et obtenues dans le cadre d'une conversation téléphonique avec un représentant de la Société canadienne des postes, le 17 novembre 2004.
9. Cette liste typologique constitue un regroupement des données fournies pour les enregistrements (Société canadienne des postes, 2005c, pages 6 à 11).

Figure 1 Fichier Recherche d'adresses de juin 2006, selon le type d'enregistrement



Source de données : Société canadienne des postes, Juin 2001, Fichier Recherche d'adresses.

Les secteurs de service des enregistrements de CPTypes 1 et 2 sont définis en fonction des adresses municipales. Les adresses municipales (de voirie) sont essentiellement celles qui se trouvent sur la façade des logements ou des immeubles. Comme on peut le voir précédemment, les CPTypes d'enregistrements correspondant aux adresses municipales constituaient 98 % des enregistrements du Fichier Recherche d'adresses en juin 2006. La plupart des côtés d'îlots urbains et de nombreuses rues résidentielles dans les régions urbaines ne comportent généralement qu'un code postal correspondant à l'adresse municipale. Même s'il peut être difficile de trouver la collectivité de la SCP et de localiser la ou les régions géographiques normalisées pour un code postal, les secteurs de service de ces codes postaux peuvent être considérés comme complètement définis dans les données sur les adresses fournies par la Société canadienne des postes. Dans ce cas, le géocodage est simple *si* le nom de rue et les données sur les adresses peuvent être appariés aux données de la même date de référence dans le cadre géographique de Statistique Canada.

Aucune donnée sur les adresses municipales des secteurs de service n'est fournie pour les CPTypes d'enregistrements 3, 4 et 5. Les secteurs de service de ces codes postaux ne sont pas directement définis dans le Fichier Recherche d'adresses. Dans le cas du CPType 3, le Fichier Recherche d'adresses fournit les codes postaux de cases postales ou de tranches de numéros de cases postales à l'intérieur d'une installation de livraison. Tous les codes postaux peuvent être couplés avec l'installation qui les dessert[10]. On peut au moins coder les codes postaux des enregistrements de CPTypes 3, 4 et 5 en fonction de la SDR où l'installation de livraison est située (si l'adresse de l'installation de livraison peut être géocodée). Dans le cas des enregistrements de CPType 4, il est parfois possible d'obtenir les noms des itinéraires desservis. Il s'agit notamment de la livraison selon des routes bien définies dans des régions rurales établies, ainsi que des livraisons à des parcs industriels et à des groupes de boîtes postales.

10. Les données du Fichier Recherche d'adresses peuvent servir à trouver l'installation de livraison dans le Fichier sur les chefs de ménage. L'adresse municipale de l'installation de livraison est souvent (mais pas toujours) disponible dans le Fichier sur les chefs de ménage. Les emplacements des installations de livraison sont aussi disponibles auprès de la Société canadienne des postes.

Pour les enregistrements de CPType 5, les seuls renseignements disponibles sont le nom de collectivité de la SCP du secteur desservi. Dans tous les cas, le secteur de service n'est pas nécessairement celui où l'installation de livraison est située, et le secteur de service n'est pas défini en fonction des tranches d'adresses municipales. En résumé, les codes postaux de CPTypes 3, 4 ou 5 (contrairement aux CPTypes 1 et 2) ne comprennent pas d'empreinte géographique précise permettant une association directe et complète avec les fichiers géographiques de Statistique Canada. D'autres sources de données décrivant les modèles d'utilisation doivent être examinées pour décider de la façon de répartir les secteurs de service pour ces codes postaux et de géocoder ces codes postaux à partir des données disponibles.

Du fait de l'urbanisation accrue et du plus grand nombre d'adresses municipales, la proportion d'enregistrements de CPType 2 a augmenté légèrement. Le tableau ci-après illustre les changements au fil du temps.

Tableau 2 Changements dans les combinaisons de type de codes postaux dans le Fichier Recherche d'adresses

CPType	Juin 2001				Juin 2006			
	Secteur de livraison urbains		Secteur de livraison rurale		Secteur de livraison urbains		Secteur de livraison rurale	
	nombre	%	nombre	%	nombre	%	nombre	%
1	739 624	98,2	0	0,0	786 435	98,2	0	0
2	0	0,0	0	0,0	0	0	0	0
3	12 039	1,6	817	15,7	12 865	1,6	841	16,1
4	250	0,0	156	3,0	249	0	136	2,6
5	626	0,1	615	11,8	647	0,1	457	8,8
1, 3	46	0,0	0	0,0	46	0	0	0
2, 4	864	0,1	290	5,6	760	0,1	443	8,5
2, 4, 5	0	0,0	106	2,0	0	0	161	3,1
2, 3, 5	0	0,0	11	0,2	0	0	39	0,8
2, 3, 4	0	0,0	88	1,7	0	0	156	3
2, 3, 4, 5	0	0,0	482	9,3	0	0	677	13
2, 5	0	0,0	6	0,1	0	0	17	0,3
3, 5	0	0,0	1 203	23,1	0	0	1 259	24,2
3, 4	0	0,0	288	5,5	0	0	212	4,1
3, 4, 5	0	0,0	864	16,6	0	0	631	12,1
4, 5	0	0,0	283	5,4	0	0	185	3,6
Total	**753 449**	**100**	**5 209**	**100,0**	**801 002**	**100**	**5 214**	**100**

0 zéro absolu ou valeur arrondie à zéro

Note : Les unités de données du tableau sont des enregistrements regroupés selon des combinaisons uniques de types de codes postaux.

Sources de données : Société canadienne des postes, Juin 2001 et Juin 2006, Fichier Recherche d'adresses.
Source du tableau : Statistique Canada, 2007, Mise en correspondance des codes postaux et des régions géographiques, numéro 92F0138MWF2007001 au catalogue.

Même si le pourcentage de chaque CPType de code postal urbain est demeuré constant, le pourcentage de codes postaux comprenant CPtype 2 a connu une légère hausse dans les secteurs de livraison rurale. La légère augmentation de la proportion de codes postaux comportant des enregistrements de tranches d'adresses municipales pourrait entraîner une légère hausse de la proportion de codes postaux pouvant être géocodés selon une tranche d'adresses municipales.

Comme on peut le voir, un code postal peut comporter plus d'un CPType d'enregistrement. Cela se produit généralement dans les régions considérées comme « rurales » par la Société canadienne des postes. Les enregistrements de CPType 2 ne viennent jamais seuls, et s'accompagnent toujours d'enregistrements de CPTypes 3, 4 et 5. Au fil du temps, la proportion

de codes postaux comportant des enregistrements de CPType 2 continue d'augmenter. Cela pourrait constituer une indication de l'augmentation du nombre d'adresses municipales, particulièrement dans les endroits où le service était associé à un itinéraire (CPType 4). Par exemple, le code postal K7V3Z8 est associé à un code postal se trouvant sur un itinéraire de distribution (CPType 4). Toutefois, il contient également des enregistrements de CPType 2 auxquels des adresses municipales sont liées. Les régions très rurales ont tendance à comporter uniquement enregistrements de CPTypes 5 et/ou 3.

Les enregistrements de CPType 1 se retrouvent dans des régions urbaines et ont tendance à être isolés, sauf dans quelques cas où ils sont liés à des enregistrements de CPType 3. Un exemple de ce genre de situation est le code postal K1A 9Z6.[11] Il est lié à une case postale et à un ensemble d'adresses municipales. L'enregistrement d'adresse municipale (CPType 1) permet de compléter l'information grâce à l'adresse de l'installation de livraison (CPType 3) pour le même code postal[12]. Effectivement, ces codes postaux comportent deux CPTypes, mais se rapportent à un seul immeuble (ou entreprise) qui reçoit du courrier.

Le secteur de service d'un code postal peut être considéré comme correspondant à un secteur où des personnes ont accès au courrier grâce à ce code postal. Dans le cas des codes postaux de CPTypes 1 et 2, il s'agit de logements et d'immeubles desservis par la poste selon leur adresse municipale. Dans le cas des codes postaux de CPType 4, il s'agit de tous les secteurs qui sont desservis par un itinéraire de distribution du courrier. Dans le cas des codes postaux de CPTypes 3 et 5, il s'agit de logements et d'immeubles dont les occupants ont généralement accès au service de livraison postale par l'entremise d'une installation de livraison. Le secteur de service d'un code postal individuel comprend chacun de ces types d'enregistrements. L'examen des combinaisons de types d'enregistrements liés à chaque code postal peut aider à déterminer si et comment la couverture de ces secteurs de service se chevauche. Cela permet de décider si certains enregistrements sont plus représentatifs de l'empreinte géographique du code postal.

Certains types d'enregistrements peuvent être considérés comme plus représentatifs du secteur de service définitif lié à ce code postal. Par exemple, le CPType 1 est plus représentatif d'un immeuble / logement particulier desservi que le CPType 5. Le CPType 4 est représentatif d'itinéraires particuliers desservis, tandis que le CPType 5 n'est pas davantage lié à un emplacement particulier qu'à un autre à l'intérieur de la collectivité de la SCP (sur la base des données sur les adresses disponibles). Comme il est indiqué précédemment dans la présente section, lorsqu'un enregistrement de type 1 et un enregistrement de CPType 3 correspondent au même code postal, l'enregistrement de CPType 1 comprend l'adresse des utilisateurs du service postal. Cela laisse supposer que les enregistrements les plus représentatifs de la couverture des codes postaux seraient les suivants :

- CPType 1 plutôt que CPType 3, lorsque les deux sont présents dans le code postal;
- CPType 2 ou CPType 4, plutôt que CPType 3 ou CPType 5, lorsque les deux sont présents dans le code postal.

Les CPTypes 2 et 4 peuvent couvrir le même secteur. Le secteur desservi par un itinéraire de distribution (CPType 4) peut être assez précis, même s'il ne peut être localisé sur la base des données sur les adresses du Fichier Recherche d'adresses. Ces données ne sont pas suffisantes pour décider lequel de ces enregistrements (CPType 2 ou 4) est le plus représentatif. Toutefois, le CPType 2 peut avoir préséance sur le CPType 4 dans le processus de géocodage, comme étant plus représentatif, étant donné qu'il est probable qu'il sera mieux géocodé (compte tenu des données particulières sur les adresses), et qu'il représente aussi probablement une partie de la couverture des enregistrements de CPType 4, par suite des progrès réalisés au chapitre des adresses municipales. Essentiellement, lorsque plus d'un type d'enregistrement est

11. Ce code postal, qui comporte des enregistrements de CPTypes 1 et 3 se retrouve dans la version de décembre 2004 du Fichier Recherche d'adresses de la Société canadienne des postes. Il a aussi été confirmé dans l'outil de recherche sur Internet consulté le 30 mai 2005, à l'adresse http://www.postescanada.ca/.
12. Ces renseignements sont fondés sur une conversation téléphonique avec un représentant de la Société canadienne des postes, le 17 novembre 2004.

lié à un code postal, le type d'enregistrement peut être utilisé comme un des critères pour choisir l'indicateur le plus représentatif de l'empreinte spatiale de ce code postal.

Une fois que la collectivité de la SCP a été appariée à une SDR, les enregistrements de CPTypes 1 et 2 peuvent être géocodés sur la base des données sur les adresses municipales, mais le manque de données sur les adresses municipales pour les types 3, 4 et 5 rend leur codage presque impossible à un niveau plus précis que la SDR. Étant donné que le but visé est de mettre en correspondance l'ensemble des empreintes de codes postaux, de plus amples renseignements sont nécessaires dans les cas où les codes postaux comportent des enregistrements de CPTypes 3, 4 et 5. Les types d'enregistrements peuvent aussi être utilisés comme facteurs pour déterminer le couplage le plus représentatif de codes postaux et de régions géographiques normalisées. En conclusion, le type d'enregistrement est essentiel pour déterminer la nature de l'empreinte du secteur de service du code postal.

Enregistrements des codes postaux et couverture de la région géographique

Il se peut qu'un code postal qui correspond spécifiquement à des tranches d'adresses municipales couvre plus d'un côté d'îlot ou îlot. Par conséquent, lorsqu'un code postal du Fichier Recherche d'adresses est lié à plus d'un enregistrement, chacun comportant sa propre tranche d'adresses, cela peut légitimement mener à la couverture de plus d'un côté d'îlot (ou îlot). Les tableaux ci-dessous montrent le nombre d'enregistrements du Fichier Recherche d'adresses liés à chaque combinaison de type d'enregistrement et de code postal.

Tableau 3 Nombre d'enregistrements selon le type d'enregistrement dans le Fichier Recherche d'adresses

CPType	Enregistrements					Tous les enregistrements d'un code postal particulier			
	Total	Minimum	Maximum	Moyen	Écart-type	Juste 1	2 to 5	6 to 50	50 et plus
	nombre					pourcentage			
Juin 2001									
1	912 835	1	78	1,23	0,94	72,56	21,20	6,20	0,04
2	58 414	1	512	31,63	41,11	0,12	1,18	47,02	51,68
3	16 718	1	34	1,06	0,49	91,59	7,36	1,05	0,00
4	4 996	1	10	1,36	0,84	57,77	40,23	2,00	0,00
5	4 196	1	1	1,00	0,00	100,00	0,00	0,00	0,00
Total	**997 159**	...	...	...	...	**68,68**	**19,80**	**8,45**	**3,07**
Juin 2006									
1	1 012 077	1	86	1,29	1,05	67,89	24,80	7,25	0,07
2	94 979	1	507	42,16	50,03	0,10	0,57	34,62	64,71
3	17 763	1	34	1,06	0,55	90,85	7,53	1,62	0,00
4	4 922	1	8	1,36	0,84	57,72	40,41	1,87	0,00
5	4 073	1	1	1,00	0,00	100,00	0,00	0,00	0,00
Total	**1 133 814**	...	...	...	...	**62,65**	**22,47**	**9,40**	**5,48**

... n'ayant pas lieu de figurer

Sources de données : Société canadienne des postes, Juin 2001 et Juin 2006, Fichier Recherche d'adresses.
Source du tableau : Statistique Canada, 2007, Mise en correspondance des codes postaux et des régions géographiques, numéro 92F0138MWF2007001 au catalogue.

L'examen des tableaux qui précèdent montre que même si les enregistrements de type 1 sont liés à une tranche d'adresses environ 70 % du temps, ils couvrent parfois plus de 50 tranches d'adresses (potentiellement 50 côtés d'îlot ou îlots). Selon notre expérience du géocodage des enregistrements de CPTypes 1 et 2, lorsqu'un code postal comprend plus d'un enregistrement de tranche d'adresses municipales, la couverture porte sur des côtés d'îlot consécutifs ou adjacents. Les enregistrements de CPType 2 semblent comporter le nombre le plus important de tranches d'adresses définies. Les enregistrements de CPType 4 comportent des itinéraires et ceux-ci peuvent couvrir de grands secteurs, même si les enregistrements de ce type ne sont pas très nombreux dans le Fichier Recherche d'adresses. Les enregistrements de CPType 3 sont généralement liés à une ou plusieurs cases postales d'une installation de livraison, et un enregistrement seulement est lié à l'installation. De même, le CPType 5 a trait à la livraison qui est disponible à partir d'une installation et peut comporter un seul enregistrement. Il convient de mentionner ici que pour les enregistrements de CPType 1, et particulièrement ceux de CPType 2, il se peut qu'aucun enregistrement géocodé (côté d'îlot ou îlot) ne soit plus représentatif qu'un autre du code postal, du point de vue du service fourni.

1.4 Fichiers de localisation des installations de livraison

Les codes postaux de CPTypes 3 et 5 sont liés à la cueillette du courrier dans une installation de livraison (plutôt qu'à sa livraison à une adresse municipale). Les adresses municipales liées à l'installation de livraison sont disponibles dans le Fichier sur les chefs de ménage. Les emplacements de ces installations de livraison ont aussi été obtenus récemment auprès de la SCP par Statistique Canada. Les emplacements ne correspondent pas toujours aux coordonnées du Fichier du réseau routier de Statistique Canada. Cela signifie que lorsque des données sur les adresses sont disponibles et peuvent être appariées aux données figurant dans le Fichier du réseau routier, il est plus approprié de procéder au géocodage à partir de celui-ci, afin que l'emplacement de l'installation par rapport aux données du réseau routier soit le plus précis possible. Lorsque des données sur les adresses ne sont pas disponibles, l'adresse de l'emplacement du bureau de poste peut être confirmée auprès de la Société canadienne des postes. Statistique Canada tiendra à jour une liste des adresses d'installations de livraison et des îlots de diffusion connexes, afin de permettre le géocodage des codes postaux. Dans la section 2.0, on tente de déterminer si certains codes postaux devraient être liés à l'installation de livraison seulement.

1.5 Type d'enregistrement et type de mode de livraison à partir du Fichier sur les modes de livraison

Le type de mode de livraison (TML) comprend des détails sur le service de livraison offert[13]. Il sert à calculer les taux pour les services postaux (Société canadienne des postes, 2005c, page 2). La description et la typologie de la SCP sont les suivantes :

A = Livraison à une adresse d'un côté d'îlot
B = Livraison à un immeuble d'appartements
E = Livraison à un immeuble commercial
G = Livraison à un gros destinataire de courrier
H = Livraison par route rurale
J = Poste restante
K = Livraison à une case postale (pas une boîte postale communautaire)
M = Livraison à un gros destinataire de courrier desservi par une case postale
T = Livraison par service suburbain
X = Livraison par itinéraire motorisé
Z = Code postal retiré (aucune autre livraison à ce code)

Le type de mode de livraison peut être utile pour illustrer si le géocodage s'applique à un immeuble particulier, soit peut-être un « sous- » côté d'îlot. Comme le montre le tableau ci-après, dans quelques cas, les enregistrements de codes postaux peuvent être liés à un immeuble particulier.

Tableau 4 Enregistrements de type 1 à partir du Fichier Recherche d'adresses de juin 2006 et TML tiré du Fichier sur les modes de livraison de juin 2006

Type	TML	Nombre d'enregistrements	Pourcentage d'enregistrements
		nombre	%
1	A	972 547	85,78
1	B	20 760	1,83
1	E	10 097	0,89
1	G	8 673	0,76

Sources de données : Société canadienne des postes, Juin 2006, Fichier Recherche d'adresses et Société canadienne des postes, Juin 2006, Données sur les modes de livraison.
Source du tableau : Statistique Canada, 2007, Mise en correspondance des codes postaux et des régions géographiques, numéro 92F0138MWF2007001 au catalogue.

Comme on peut le voir, les combinaisons les plus fréquentes de TML et de CPType sont liées à des tranches d'adresses bien définies (TML = « A »). Les autres CPType 1 sont liés à des immeubles d'appartements (TML = « B ») et à différents types d'entreprises (TML = « E » et « G »). La majeure partie du géocodage des enregistrements de CPType 1 peut simplement être effectuée à partir des données sur les adresses. Toutefois, il est particulièrement important que les immeubles comportant des populations importantes soient géocodés de façon appropriée. Par exemple, un immeuble peut correspondre à un établissement de santé qui représente un îlot,

13. Le TML ne doit pas être confondu avec le CPType du FCCP. Le TML a davantage à voir avec la **façon** dont le courrier est livré. Le type d'enregistrement du code postal a préséance du point de vue de l'**endroit** où le courrier est livré. Un cas type est celui d'un gros destinataire de courrier dont le code postal comporte des enregistrements de CPTypes 3 et 1. Un envoi dont l'adresse est formatée comme un enregistrement de CPType 1 (adresse municipale) continuera d'être envoyé à la case postale correspondant au code postal, à condition que le code postal soit correct et puisse être lu à la machine. (Autrement, la machine tentera de déchiffrer le code postal sur la base des données de l'adresse.) Ce système empêche la livraison de quantités importantes de courrier par un facteur. Le TML est utilisé par les gros envoyeurs de courrier pour obtenir des remises, par suite du tri préliminaire du courrier. (Ces données sont fondées sur une conversation avec un représentant de la SCP, le 27 novembre 2004.)

une aire de diffusion ou une subdivision de recensement *(SDR)* complets. Dans de rares cas où une SDR complète est représentée dans un établissement, sa population peut être équivalente à l'ensemble de la population de la SDR. Pour les codes postaux de CPType 1, le TML permet de faire une distinction entre le type d'immeuble (logement, appartement, gros destinataire de courrier, etc.) qui est représenté par le code postal. Le TML ne comprend pas plus de renseignements que le type d'enregistrement pour localiser les types de codes postaux d'après les données sur les adresses, bien qu'il donne de l'information sur le genre d'immeuble qui reçoit le courrier.

1.6 Considérations relatives au géocodage fondé sur les données de la SCP

Ces considérations relatives au géocodage font suite à l'examen des données de la SCP de la présente section. Ces considérations, de pair avec les conclusions de la section suivante sur la façon dont les codes postaux sont utilisés et déclarés, permettront d'élaborer les règles du géocodage.

Les considérations relatives à l'élaboration des règles de géocodage sont les suivantes[o]:

- Le nom de collectivité de la SCP et la première lettre du code postal constituent parfois, mais pas toujours, une façon de localiser la SDR pour le géocodage. Même si, en théorie, plus de 97 % des enregistrements du Fichier Recherche d'adresses comprennent des données sur les adresses municipales, moins de 40 % des noms uniques de collectivités de la SCP peuvent être appariés à la structure de la SDR. Cela fait ressortir dès le départ la nécessité de données supplémentaires pour confirmer et compléter le géocodage.
- Dans le cas des CPTypes 1 et 2, le nom de collectivité de la SCP combiné aux données sur les adresses municipales permettent de déterminer uniquement les adresses municipales (représentant les logements et les immeubles) desservies par le code postal, **à condition** que le secteur qui est représenté par le nom de collectivité de la SCP puisse être trouvé.
- Dans le cas des CPTypes 3, 4 et 5, le secteur de service n'est pas défini au moyen des adresses municipales. Des données supplémentaires sur la façon dont ces codes postaux sont utilisés sont nécessaires pour les géocoder.
- Les combinaisons de types d'enregistrements pour un code postal peuvent servir à déterminer les enregistrements les plus représentatifs ou indicatifs pour ce code postal. Toutefois, il peut y avoir plus d'un enregistrement de tranche d'adresses lié à un type d'enregistrement pour un code postal, et ces codes postaux peuvent de façon assez légitime être liés à plus d'un côté d'îlot ou îlot.

En conclusion, des données supplémentaires sur les secteurs desservis par le code postal sont nécessaires pour élargir et valider le géocodage partiel qui peut être effectué à partir des fichiers de la SCP seulement.

2.0 Utilisateurs des codes postaux et géographie

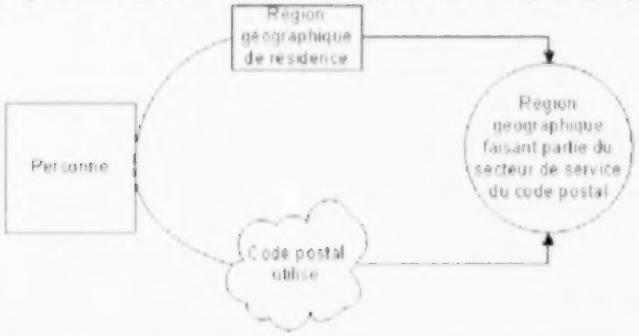
Après examen des données sur les codes postaux dans la section précédente, le défi principal subsiste : comment élaborer un secteur de recherche à partir duquel le code postal peut être géocodé. Les noms de collectivités de la SCP représentent des noms de lieux qui ne correspondent habituellement pas aux noms ou aux limites des SDR du cadre géographique de Statistique Canada. Dans la section qui précède, la discussion sur l'utilisation des fichiers de la SCP pour le géocodage a montré que des données supplémentaires sont nécessaires. Les questions de fond qui restent sans réponse sont les suivantes :

1. Étant donné que moins de 40 % des noms uniques de collectivités de la SCP correspondent à une SDR, comment pouvons-nous géocoder les adresses comprenant les 60 % qui restent de noms de collectivités de la SCP?
2. Comment le secteur de service peut-il être déterminé pour les enregistrements de types 3, 4 et 5, pour lesquels nous n'avons pas d'adresses municipales?
3. Existe-t-il des données supplémentaires qui peuvent servir à valider le géocodage fondé sur les données sur les adresses?

Les renseignements supplémentaires dont il est question ici sont des renseignements fondés sur l'utilisation déclarée du code postal. Le secteur où le code postal est utilisé est appelé secteur de service du code postal. Les secteurs de service, sur la base des codes postaux déclarés dans le Recensement de la population, fournissent les renseignements supplémentaires nécessaires pour assurer et confirmer le géocodage. Cela s'ajoute à la recherche de la SDR, sur la base du nom de collectivité de la SCP. Des améliorations à ce processus sont examinées à l'annexe C. La présente section met l'accent sur la façon dont les personnes ou les ménages utilisent les codes postaux, et comment ceux-ci viennent compléter et confirmer les données sur les adresses de la SCP.

Les utilisateurs des codes postaux sont des personnes et des entreprises. Toutefois, le rapport entre ces unités statistiques et les codes postaux est indirect. Comme il est indiqué dans la section précédente, les codes postaux ne sont pas attribués à des personnes ou des entreprises, mais bien à des immeubles et des logements[14]. L'exception à cette règle est le fait que moins de 2 % des enregistrements d'adresses de codes postaux ont un type de mode de livraison « G » ou « M », lesquels sont associés à un gros destinataire de courrier. Une source de données est aussi nécessaire pour déterminer le secteur de service de cette façon. Dans le cas des personnes, les rapports suivants doivent exister entre les sources de données :

Figure 2 Code postal utilisé en rapport avec la région géographique de résidence



14. La grande majorité des enregistrements du Fichier Recherche d'adresses sont de CPTypes 1 et 2 et ont trait à la livraison à un logement ou un immeuble particulier. Les autres types d'enregistrements ont trait aux services disponibles à partir de l'installation de livraison.

Essentiellement, le rapport de la personne avec le code postal et la région géographique normalisée peut servir à faire le lien entre ces derniers.

Même si une telle source de données est disponible pour le code postal, la question subsiste quant à la région géographique normalisée liée à ce code postal. Par exemple, il y aurait un risque que la région géographique normalisée attribuée dans la source des données ne corresponde pas à celle du domicile de la personne, mais soit liée à la région géographique normalisée de travail. Ces risques doivent être équilibrés du point de vue de la façon dont les personnes et les entreprises utilisent les codes postaux.

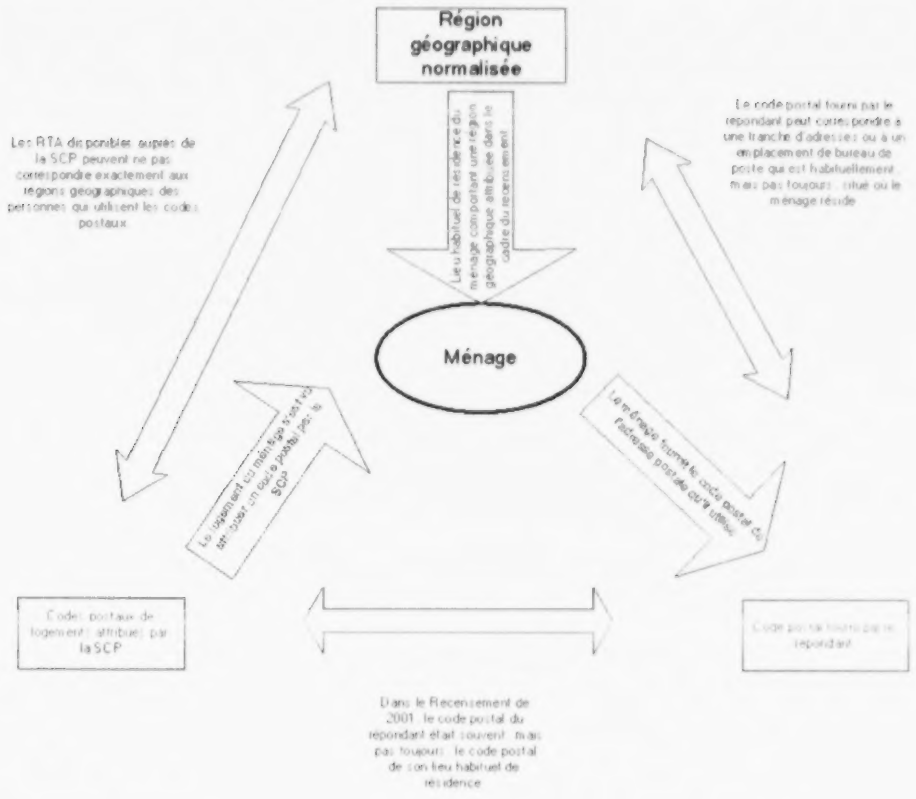
2.1 Personnes et codes postaux déclarés dans le recensement

Le Recensement de la population est la seule source comprenant un code postal déclaré et la région géographique normalisée pour toutes les personnes au Canada. Les ménages au Canada doivent déclarer (ou confirmer) leur code postal dans le formulaire de recensement. Ce code postal de répondant correspond généralement au code postal de la résidence du ménage. Il peut être utilisé pour estimer le secteur de service d'un code postal.

Un ménage est généralement défini comme étant constitué d'une personne ou d'un groupe de personnes qui résident ensemble dans un logement ou qui occupent un logement.

Le recensement est l'enquête permettant de recueillir les données les plus complètes sur les codes postaux des ménages pour le Canada. Le code postal fourni par le répondant dans le cadre du recensement et le code postal attribué au logement de la personne par la SCP sont généralement les mêmes. Ces données peuvent servir à déterminer la région géographique normalisée liée au code postal déclaré dans le cadre du recensement. Le diagramme ci-après montre comment ces éléments sont liés.

Figure 3 Utilisation des codes postaux par les ménages



Dans le cadre du recensement, on demande aux répondants de fournir le code postal de leur résidence. Étant donné que ces renseignements sont disponibles au niveau du ménage et au niveau de l'îlot de diffusion, pourquoi ne pas simplement attribuer des codes postaux aux îlots de diffusion (et autres régions géographiques normalisées), sur la base du recensement seulement? Les régions géographiques normalisées servent généralement de base à l'analyse, et toute extraction de données à partir du code postal peut être effectuée à partir de ces régions géographiques normalisées. Toutefois, si les régions géographiques normalisées doivent être utilisées comme moyen pour coupler de façon fiable les codes postaux et d'autres données, les données du recensement doivent être complétées par le géocodage pour les raisons suivantes :

1. Les codes postaux changent chaque mois, et des centaines de codes sont créés, modifiés et supprimés. Les données du recensement sont disponibles uniquement tous les cinq ans, et de nombreux mois après le jour du recensement. L'opération de géocodage permet le géocodage opportun de la plupart des codes postaux courants.

2. À l'occasion, plus d'un code postal peut être utilisé par une personne ou une entreprise. Il se peut qu'une personne reçoive du courrier à son lieu habituel de résidence, qu'une personne occupée qui voyage utilise son adresse au travail et y fasse acheminer son courrier, qu'un étudiant reçoive du courrier à l'adresse de ses parents, ou encore que certaines personnes préfèrent recevoir leur courrier au bureau de poste. Ce sont tous là des exemples d'adresses différentes et de codes postaux différents qui peuvent être utilisés par la même personne. De même, une entreprise peut utiliser un code postal lié à

l'immeuble ou au côté d'îlot où est situé cet immeuble, à une case postale, ou au code postal de l'adresse postale de son siège social.

3. La variable de réponse au code postal doit être vérifiée parce qu'il peut y avoir des variations dans la réponse au code postal (comme il est indiqué précédemment). Le codage des codes postaux en fonction d'un Fichier du réseau routier du cadre géographique de Statistique Canada permet de confirmer que la région géographique normalisée est celle de l'empreinte du code postal.

Les codes postaux déclarés dans le cadre du recensement peuvent servir à mettre en correspondance les secteurs de service, tant pour le code postal que pour la RTA. Ainsi, on peut obtenir des données de meilleure qualité en combinant les données du recensement et celles disponibles auprès de la SCP.

2.2 Identification du secteur de recherche pour le géocodage

Étant donné que moins de 40 % des noms uniques de collectivités de la SCP correspondent à une SDR, comment pouvons-nous géocoder les adresses comprenant les 60 % qui restent de noms de collectivités de la SCP? On peut localiser les noms de rue et les adresses pour le géocodage à partir des données du recensement.

Un essai a été mené à l'automne 2005, afin de mesurer le taux de géocodage fondé sur les secteurs de service de RTA. Étant donné que les RTA sont plus stables que les codes postaux au fil du temps, des secteurs de recherche ont été créés à partir des RTA des codes postaux valides déclarés dans le Recensement de 2001. Pour l'essai de géocodage des codes postaux de juin 2005, des *secteurs de service de RTA* ont été créés à partir des codes postaux valides en mai 2001. Des îlots de diffusion ont été attribués à une RTA uniquement si un code postal valide comportant cette RTA était déclaré au moins quatre fois dans cet îlot de diffusion[15]. Le géocodage des codes postaux a été effectué à l'intérieur du secteur de service de la RTA pour la RTA de ce code postal. Essentiellement, les secteurs de service de RTA ont servi de secteurs de recherche pour le géocodage. En dépit de ces critères conservateurs, plus de 88 % des enregistrements comportant une adresse municipale (de CPTypes 1 et 2) ont pu être géocodés, la plupart, au niveau le plus détaillé du côté d'îlot.

L'utilisation de ces secteurs de recherche semble être une façon productive d'effectuer le géocodage, lorsque le secteur de service du code postal est désigné au moyen d'adresses municipales (CPTypes 1 et 2) et que les adresses municipales et les données du réseau routier connexes sont disponibles dans le Fichier du réseau routier.

2.3 Géocodage des codes postaux directement à partir des données du recensement

Une méthode est nécessaire pour le géocodage des codes postaux qui ne sont pas liés à des adresses municipales (CPTypes 3, 4 ou 5) ou dont les adresses municipales ne se trouvent pas dans le Fichier du réseau routier. Est-ce que le recensement fournit des renseignements suffisants pour les géocoder?

15. Lorsque les codes postaux d'une autre province ont été déclarés à l'intérieur de l'îlot de diffusion, cela n'a pas été pris en compte dans le géocodage. La méthode de géocodage est expliquée de façon plus détaillée dans la section suivante.

Les données sur le code postal déclarées dans le cadre du Recensement de 2001 peuvent être examinées, afin de déterminer les types de codes postaux qui ont été utilisés. Ces réponses devraient être évaluées avec précaution, certains codes postaux comportant plusieurs types d'enregistrements, et les répondants n'indiquant pas le type d'enregistrement de leur code postal. Les codes postaux sont par conséquent examinés ci-après dans le contexte des combinaisons de types d'enregistrements. Le Fichier Recherche d'adresses de juin 2001 comportait 758 658 codes postaux qui n'ont pas été retirés. Parmi les 673 242 codes postaux valides déclarés dans le cadre du recensement, 673 179 ont été trouvés dans le Fichier Recherche d'adresses de juin 2001.

Un examen détaillé des codes postaux qui n'ont ***pas*** été déclarés a révélé ce qui suit.

Tableau 5 Codes postaux qui n'ont *pas* été déclarés dans le Recensement de 2001

	Livraison urbaine			Livraison rurale		
	Juin 2001	Recencement de 2001		Juin 2001	Recencement de 2001	
CPType	Fichier Recherche d'adresses	Non déclarés	% de CP non déclarés	Fichier Recherche d'adresses	Non déclarés	% de CP non déclarés
	nombre	nombre	%	nombre	nombre	%
1	739 624	76 581	10,35	0	0	0,00
2	0	0	0,00	0	0	0,00
3	12 039	6 572	54,59	817	6	0,73
4	250	35	14,00	156	2	1,28
5	626	331	52,88	615	8	1,30
1, 3	46	3	6,52	0	0	0,00
2, 4	864	12	1,39	290	0	0,00
2, 4, 5	0	0	0	106	0	0,00
2, 3, 5	0	0	0	11	0	0,00
2, 3, 4	0	0	0	88	0	0,00
2, 3, 4, 5	0	0	0	482	0	0,00
2, 5	0	0	0	6	0	0,00
3, 5	0	0	0	1 203	2	0,17
3, 4	0	0	0	288	0	0,00
3, 4, 5	0	0	0	864	1	0,12
4, 5	0	0	0	283	0	0,00
Total par CP	**753 449**	**83 534**	**11,09**	**5,209**	**19**	**0,36**

0 zéro absolu ou valeur arrondie à zéro

Sources de données : Société canadienne des postes, Juin 2001, Fichier Recherche d'adresses et Statistique Canada, Recensement de la population de 2001

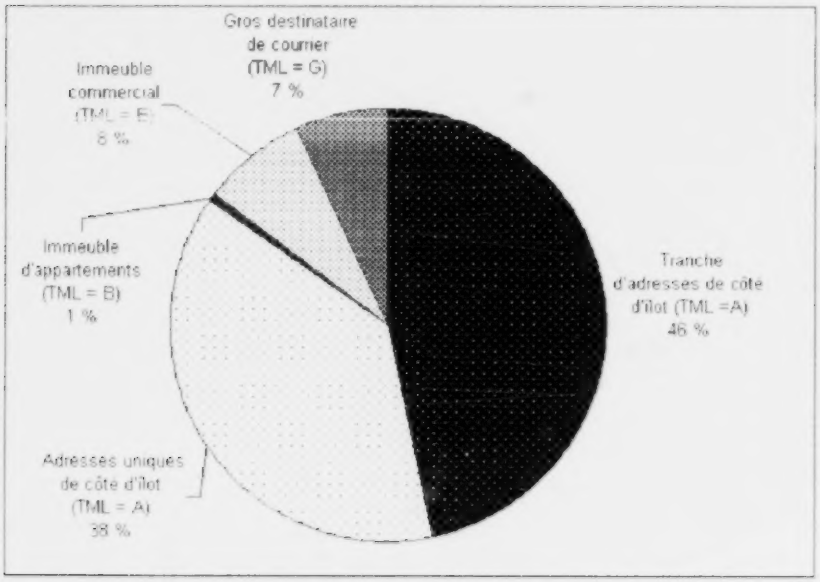
Source du tableau : Statistique Canada, 2007, Mise en correspondance des codes postaux et des régions géographiques, numéro 92F0138MWF2007001 au catalogue.

Il semble que presque tous les codes postaux de livraison rurale sont déclarés. Seulement 19 codes postaux de livraison rurale n'ont pas été déclarés du tout. Cela représente moins de 1 % de tous les codes postaux ruraux. Étant donné que le codage des enregistrements de CPTypes 4 et 5 est particulièrement difficile, l'utilisation des réponses du recensement semble être une approche efficace pour géocoder ces codes dans les régions rurales. Une comparaison des résultats déclarés du Recensement de 2001 et de ceux du Recensement de 2006 devrait nous permettre de déterminer si ces codes postaux sont stables au fil du temps et si, par

conséquent, les données du recensement sont suffisantes pour géocoder les codes postaux actifs au moment du recensement.
Un moins grand nombre de codes postaux de livraison urbaine ont été déclarés (en comparaison avec la livraison rurale). Les codes postaux des entreprises ne sont pas déclarés dans le recensement pour des raisons évidentes. Cela semble être le cas pour environ 50 % des codes postaux de CPType 3. Les codes postaux de CPType 5 ne sont peut-être pas utilisés autant dans les secteurs de livraison urbaine, étant donné particulièrement que la plupart des personnes se font livrer le courrier directement à leur résidence. Même si seulement 10 % des codes postaux de type 1 n'ont pas été déclarés, leur nombre est quand même important (76 581). Ces codes postaux ne sont pas considérés comme des codes postaux d'entreprises selon le type d'enregistrement. Toutefois, après un examen plus approfondi, il semble s'agir de codes postaux liés à des immeubles non résidentiels. Par exemple, K2E 7L6, K2E 7J5 et K2E 1B6 sont associés à des adresses municipales le long de la promenade Colonnade à Ottawa. Celle-ci est zonée comme une « zone industrielle générale »[16]. Les sous-ensembles des codes postaux de CPType 1 qui n'ont pas été déclarés ont été classés par la suite du point de vue de leur TML, puis vérifiés, afin de déterminer s'ils correspondaient à une livraison à une adresse unique.

La ventilation des codes postaux de CPType 1 qui n'ont pas été déclarés figure ci-après.

Figure 4 Codes postaux de CPType 1 non déclarés



Sources de données : Société canadienne des postes, Juin 2001, Données sur les modes de livraison et Statistique Canada, Recensement de la population de 2001.

Ensemble, un peu plus de la moitié de ces codes postaux correspondent à une livraison à une adresse unique[17]. Étant donné que plus de la moitié de ces codes postaux sont liés à une adresse unique, et certains d'entre eux sont clairement désignés comme des adresses commerciales par la SCP, cela laisse supposer que les codes postaux urbains de CPType 1 pourraient en fait être principalement des codes postaux d'entreprises ou d'établissements.

16. À partir des cartes et des données du site Web de la ville d'Ottawa (*Ottawa.ca*), consulté le 5 juillet 2006.
17. Cela comprend tous les codes postaux du graphique précédent dont les TML sont E, B et G, ainsi que les codes postaux dont le TML est A et qui correspondent à une livraison à une adresse uniquement seulement.

L'utilisation du code postal déclaré dans le recensement constitue probablement un outil approprié pour déterminer les secteurs de service de RTA, tout comme celle des codes postaux liés à des ménages (plutôt qu'à des entreprises ou des établissements). Généralement, les codes postaux de ménages sont déclarés dans le recensement et les secteurs de service créés à partir de la RTA déclarée devraient être suffisants pour le géocodage de la majorité des codes postaux[18].

2.4 Recommandations pour le géocodage fondé sur les réponses du recensement

Les données du recensement constituent une bonne source de géocodage, dans les cas où les codes postaux sont utilisés au moment du recensement. Elles peuvent aussi servir à créer les secteurs de recherche pour le géocodage des codes postaux d'adresses municipales. Ces données de base peuvent aider à attribuer les codes postaux aux îlots de diffusion, *aires de diffusion* et SDR.

L'utilisation du code postal peut changer au fil du temps parce que la Société canadienne des postes modifie les codes postaux et que les utilisateurs peuvent modifier leurs codes postaux déclarés. Par ailleurs, il n'est pas possible d'obtenir un code postal pour chaque ménage dans le cadre du recensement, parce que les réponses sont parfois illisibles, invalides ou laissées en blanc. Toutefois, le code postal figurant dans le recensement, particulièrement s'il est fréquemment déclaré à l'intérieur d'un îlot de diffusion, peut être utilisé efficacement pour définir la RTA du code postal utilisé selon l'îlot. Même si les codes postaux ont tendance à changer un peu chaque mois, le secteur de service de la RTA sera probablement plus stable sur la période de cinq ans séparant les recensements[19]. La possibilité de géocoder les données de la Société canadienne des postes pour 2005 à partir des secteurs de service créés sur la base des données du Recensement de 2001 laisse supposer que les RTA sont en fait habituellement stables. La stabilité des secteurs de service de RTA, ainsi que du rapport entre les codes postaux déclarés et les régions géographiques normalisées, devraient aussi faire l'objet d'une vérification plus rigoureuse, grâce à la comparaison des résultats du Recensement de 2001 et de ceux du Recensement de 2006.

Les codes postaux déclarés dans le Recensement de 2006 peuvent être utilisés de la façon suivante pour le géocodage.

- Pour les CPTypes 1 et 2 qui desservent des ensembles particuliers de logements ou d'immeubles, le but visé est le géocodage en fonction de la tranche d'adresses du côté d'îlot. Le secteur de service de la RTA peut être établi avec un niveau de confiance raisonnable. Les enregistrements du Fichier Recherche d'adresses peuvent être géocodés à l'intérieur de ces secteurs de RTA. Les secteurs de recherche de RTA peuvent être complétés par des secteurs de recherche créés à partir du rapport entre le nom de la SDR du recensement courant et les noms de la SDR du recencement précédent. On peut tenter le géocodage selon la région géographique normalisée la plus détaillée en cherchant le nom de rue et la tranche d'adresses dans le secteur de service.
- Les enregistrements de CPType 4 ne comprennent pas de tranches d'adresses municipales. Toutefois, la déclaration de ces codes postaux permettrait de déterminer où

18. La méthode utilisée pour le géocodage des codes postaux en fonction des régions géographiques du Recensement de 2001 est fondée sur l'appariement du nom de collectivité de la SCP et du nom de SDR. Cette méthode pose des problèmes, comme il est indiqué dans la section 1.3. Toutefois, des améliorations sont apportées à l'appariement des noms de SDR et sont abordées plus loin dans l'annexe C.
19. Lorsque des changements à grande échelle touchent à l'occasion les secteurs de RTA, ils sont signalés par la SCP.

ils sont utilisés, et les mêmes codes postaux déclarés pour ce secteur dans le cadre de plus d'un recensement confirmeraient la stabilité de l'utilisation — d'itinéraires bien définis.

- Le service est assuré à partir de l'installation de livraison pour les CPTypes 3 et 5. Au premier abord, il semble que ceux-ci devraient être géocodés uniquement en fonction de l'installation de livraison postale. Toutefois, les données du recensement montrent que dans les secteurs de livraison rurale, ces codes postaux sont couramment utilisés et sont peut-être les seuls disponibles. Dans les secteurs de livraison rurale, un code de CPType 5 peut être attribué sur la base des données du recensement. Dans de rares cas, lorsqu'il n'y a pas de codes postaux déclarés pour une région rurale habitée, il est probable que lorsqu'un code postal est utilisé, il s'agit du code postal de l'installation de livraison la plus proche.

Le géocodage directement en fonction des régions géographiques normalisées à partir des données du recensement, plutôt que par suite de la conversion des données géocodées précédemment, devrait être conforme au concept de secteurs de service de code postal et devrait permettre une couverture plus complète et une précision plus grande.

3.0 Règles de géocoage et mesures de la qualité

On peut procéder au géocodage sur la base de règles, en combinant les données du recensement et les fichiers de la SCP. Des indicateurs de qualité peuvent être élaborés pour chaque enregistrement géocodé au moyen de cette méthode, la méthode de géocodage fondée sur des règles et les données utilisées pour le géocodage pouvant servir à créer les indicateurs de qualité.

3.1 Secteurs de service et secteurs de recherche

Les secteurs de service doivent être précisés pour que le processus de géocodage puisse être entrepris. Les secteurs de service et les rapports proposés pour le géocodage sont les suivants :

1. Secteur de service de la RTA (fondé sur le recensement)
2. Secteur de service du code postal (fondé sur le recensement)
3. Fichier de correspondance de la collectivité de la SCP et de la SDR

Comme il est indiqué précédemment, le secteur de service de la RTA peut servir à la recherche et au géocodage en fonction des tranches d'adresses municipales. Dans le cas du Recensement de 2006, chaque secteur de service de la RTA engloberait toutes les AD qui comprennent au moins un îlot de diffusion où un nombre minimum de répondants ont déclaré cette RTA dans le cadre du recensement[20]. Le fait de compter un nombre minimum de répondants permet de s'assurer que la RTA est véritablement utilisée et déclarée de façon précise à l'intérieur de l'AD. Étant donné que plus d'une RTA peut être utilisée à l'intérieur d'une AD, les limites de RTA élaborées pour des AD globales se chevaucheront. Toutefois, un léger chevauchement et un surdénombrement ne devraient pas poser de problème, étant donné que ces secteurs de service de RTA ne seront pas utilisés directement pour procéder au géocodage. Les secteurs de service de RTA fourniront uniquement les secteurs de recherche pour la recherche et le géocodage des codes postaux, sur la base des données sur les adresses municipales figurant dans les fichiers du réseau routier.
Le secteur de service du code postal, contrairement au secteur de service de la RTA, sera élaboré de façon plus rigide, afin d'inclure uniquement les îlots de diffusion où ce code postal a

20. Pour la vérification des secteurs de service de RTA (mentionnée dans la section 2.2), un îlot de diffusion est attribué à un secteur de service uniquement si un minimum de quatre répondants ont déclaré des codes postaux comportant

été déclaré par un nombre minimum de répondants. Par ailleurs, afin de limiter les chevauchements, seuls les codes postaux les plus souvent déclarés dans cet îlot de diffusion seront retenus. Des critères plus rigides de délimitation sont utilisés parce que le secteur de service du code postal peut servir à géocoder directement le code postal en fonction d'un îlot de diffusion où ce code postal est utilisé.

Toutefois, on ne peut pas attribuer un secteur de recherche à tous les codes postaux sur la base des données fournies dans le Recensement de 2006[21]. Des codes postaux imputés pourraient être utilisés pour créer des secteurs de recherche plus exhaustifs, mais cette façon de faire est généralement un peu moins fiable. Il existe une autre possibilité, à savoir créer le lien entre la SDR utilisée pour le recensement et les anciens noms de la SDR, ainsi que le fichier de noms de lieux de la Division de la géographie.

La couverture de la collectivité de la SCP montre le secteur où le nom de rue est unique. (Comme il est indiqué précédemment, la SCP conserve les anciens noms de municipalités dans les cas de fusions et de modification des limites, tant que les problèmes liés aux adresses municipales en double ne sont pas résolus.) Il est difficile de définir de façon définitive les limites des collectivités de la SCP, parce que ces limites ne sont pas disponibles auprès de la SCP. Toutefois, les SDR peuvent être appariées à une collectivité de la SCP [22]. Cet appariement entre le nom de la SDR et la collectivité de la SCP peut servir à confirmer le géocodage automatisé selon le secteur de service de la RTA. Toutefois, étant donné que la correspondance est floue du point de vue du secteur couvert, elle ne servira pas de méthode principale pour créer les secteurs de recherche pour le géocodage automatisé. La correspondance de la collectivité de la SCP et de la SDR servira uniquement comme outil de vérification additionnelle du géocodage automatisé et comme méthode secondaire de géocodage.

3.2 Géocodage des codes postaux selon le CPType

Étant donné que le but visé est de géocoder les codes postaux de la façon la plus définitive possible, la priorité sera accordée au géocodage des enregistrements de codes postaux dont les secteurs de service sont définis selon des tranches d'adresses municipales. Il s'agit des codes de CPTypes 1 et 2. La couverture la plus définitive suivante est probablement celle du CPType 4, étant donné que les itinéraires desservis sont définis, même s'ils ne comportent pas d'adresses municipales ou de noms de rue qui peuvent être appariés à un Fichier du réseau routier. Enfin, les CPTypes 3 et 5, qui sont liés à une installation de livraison postale, mais qui peuvent être utilisés par les ménages de divers secteurs de la collectivité de la SCP ou même par des voyageurs ou d'autres personnes à l'extérieur de la municipalité, auront la priorité en ce qui a trait à l'affectation des régions géographiques normalisées au code postal. Cette affectation des régions géographiques normalisées dans le contexte du type de service disponible à l'endroit en question met l'accent sur le code postal le plus susceptible d'être utilisé dans cette région géographique normalisée.

---

cette RTA dans cet îlot. Le choix d'un nombre minimum de réponses pour une RTA est fondé sur l'hypothèse que la probabilité est plus faible que quatre personnes fassent la même erreur en déclarant la RTA dans le même îlot que celle d'une personne de faire l'erreur dans la même îlot. Malgré l'utilisation de ce critère conservateur, le taux de géocodage a été supérieur à 88 % pour les enregistrements d'adresses municipales de CPType 1 et 2.

21. D'après la vérification des secteurs de service de la RTA mentionnée précédemment, les secteurs de service n'ont pu être élaborés pour tous les codes postaux, ces derniers n'ayant pas tous été utilisés.
22. Une liste de correspondance entre les noms de collectivités de la SCP et les SDR a été créée pour la recherche d'adresses municipales dans le cadre du processus de géocodage manuel. Voir l'annexe C pour plus de détails sur les améliorations apportées à ce fichier de correspondance.

Les codes postaux de CPTypes 1 et 2 peuvent définitivement être géocodés grâce à l'appariement des données sur les adresses de la SCP et de celles du Fichier du réseau routier. Si l'appariement repose uniquement sur les données du réseau routier et celles sur les adresses, on pourra obtenir des appariements multiples (dans l'ensemble du Canada). Par exemple, il existe de nombreuses rues « Principale » au Canada. La recherche du nom de rue « Principale » donnera lieu à des appariements multiples. Dans le programme d'appariement automatisé, les appariements seront effectués à l'intérieur d'un secteur de recherche. Les essais effectués jusqu'à maintenant laissent supposer que le secteur de recherche de la RTA peut être utilisé comme secteur de recherche principal. La SDR de l'appariement peut aussi être confirmée par le fichier de correspondance de la collectivité de la SCP et de la SDR (voir l'annexe C sur la création d'une correspondance entre la collectivité de la SCP et la SDR). Dans la mesure du possible, les enregistrements de CPTypes 1 et 2 devraient être géocodés selon le côté d'îlot ou l'îlot de diffusion, étant donné qu'ils correspondent de toute évidence à la livraison postale à des tranches d'adresses. La précision du géocodage peut aussi être contrôlée – le nombre d'*îlots de diffusion* ou de côtés d'îlot géocodés devrait être proportionnel aux enregistrements de tranches d'adresses du Fichier Recherche d'adresses de la SCP. De cette façon, les données du recensement et les données du Fichier Recherche d'adresses de la SCP peuvent être combinées pour procéder au géocodage et en évaluer la qualité.

Diverses combinaisons de CPTypes 2, 3, 4 et 5 sont obtenues en parallèle, particulièrement dans les secteurs de livraison rurale (voir le tableau 5). Les codes postaux déclarés dans le recensement ne comprennent pas de renseignements sur le type d'enregistrement. Par conséquent, le CPType de service lié au code postal ne peut être déduit. Le Fichier Recherche d'adresses de la SCP comprend des données sur l'adresse municipale pour les codes de CPType 2, mais la couverture de chacun des autres CPTypes, pour un code postal donné, ne peut être déduite. Compte tenu de cela, le géocodage d'une partie d'un type particulier de code postal de CPType 3, 4 ou 5 peut ne pas être précis du point de vue de l'utilisation. Par le passé, les codes postaux liés à des installations de CPType 3 principalement devaient être codés (dans le FCCP) selon l'îlot de l'installation. L'approche dans ce cas consiste à poursuivre ce codage, tout en désignant l'enregistrement du FCCP comme se rapportant à l'installation. Toutefois, lorsque la poste restante (CPType 5) et la case postale (CPType 3) sont les seuls services disponibles, ils peuvent aussi être géocodés selon la région géographique normalisée, du fait qu'il s'agit du code postal le plus couramment utilisé dans ce cas. L'approche qui suit est proposée pour les codes postaux comportant toute permutation d'enregistrements de CPTypes 2, 3, 4 et 5.

1. Si le code postal comporte des enregistrements de CPType 2, à l'intérieur du secteur de recherche de la RTA, coder tous les enregistrements de CPType 2 selon les données sur les adresses du Fichier Recherche d'adresses.
2. Si les îlots de diffusion du secteur de service du code postal n'ont pas été complètement géocodés à l'étape précédente, attribuer les îlots de diffusion restants qui sont géocodés de façon incomplète (ou qui ne sont pas géocodés) et les AD de ce secteur de service de code postal le plus possible en fonction des enregistrements de CPType 4. Lorsque les codes postaux sont liés à des enregistrements de CPType 4 seulement ou à des enregistrements de CPTypes 4 et 2 seulement, *tous* les enregistrements restants à ce moment-là peuvent être attribués au CPType 4. Toutefois, lorsqu'un enregistrement de CPType 4 s'accompagne d'enregistrements de CPType 3 et/ou 5 (moins de 0,5 % des codes postaux), il se peut qu'un traitement manuel au moyen de cartes soit nécessaire pour assurer la couverture appropriée des itinéraires liés aux enregistrements de CPType 4.
3. Toutes les régions géographiques normalisées restantes du secteur de service du code postal qui n'ont pas été attribuées pourraient l'être aux enregistrement de CPType 5 au niveau de l'AD, selon le principe que ce service est accessible à partir de l'installation de livraison postale pour les ménages de ce secteur. De même, lorsqu'un service de poste restante (CPType 5) n'est pas disponible, les îlots de diffusion pourraient être codés selon l'enregistrement de CPType 3 de ce secteur du service de code postal.

Dans les AD peu peuplées, où le service postal utilisé n'est pas évident à partir des réponses fournies ou des données sur les adresses, ce service est probablement assuré à partir de l'installation postale la plus proche. Ces AD pourraient être codées selon le code postal de l'installation de livraison la plus proche où le service postal est accessible (type 5 ou 3). Le géocodage effectué de cette façon pourrait être signalé à l'aide d'un champ indiquant la source de géocodage. Le but visé est d'assurer le géocodage complet des codes postaux dans le contexte du service disponible dans chaque AD habitée.

La méthode de géocodage décrite ici repose sur le secteur de service de la RTA, le recensement servant de secteur de recherche principal. Les enregistrements de CPTypes 1 et 2 sont par la suite géocodés à l'intérieur de ce secteur, grâce à un appariement en fonction des données sur les adresses municipales. Lorsque cela n'est pas possible, le rapport entre la collectivité de la SCP et la RTA à la SDR pourra être utilisé. Le géocodage des enregistrements de PCTypes 3, 4 et 5 est plus complexe, étant donné que ceux-ci ne correspondent pas à des adresses municipales. Toutefois, les réponses du recensement peuvent être utilisées pour effectuer le codage dans les secteurs où ces codes postaux sont utilisés. Les enregistrements de PCTypes 3 et 5 continueront aussi d'être codés selon l'installation postale, ainsi que le secteur de service. Ainsi, tous les codes postaux utilisés peuvent être mis en correspondance avec des régions géographiques, sur la base des données du recensement disponibles et du Fichier du réseau routier.

3.3 Mesures de la qualité

La qualité du géocodage des codes postaux peut être améliorée et faire l'objet de meilleurs rapports, compte tenu du concept de géocodage selon la RTA et les secteurs de service de code postal. (Le contexte et les définitions des mesures de la qualité figurent à l'annexe B.)

A. Actualité

L'actualité peut être mesurée du point de vue de la différence entre les éléments suivants :

1. La date de référence des codes postaux et la date de diffusion des produits postaux.
2. La date de référence du recensement et la diffusion des produits fondés sur les codes postaux.

L'examen conceptuel dont il est question précédemment prévoit le géocodage fondé sur des règles pour les codes postaux, selon les côtés d'îlot, les îlots de diffusion et les AD du cadre géographique de Statistique Canada. Le géocodage automatisé a été mis à l'essai et peut améliorer l'efficacité du processus. Même s'il n'est pas possible d'utiliser le géocodage automatisé lorsque les données sont insuffisantes dans le cadre géographique ou qu'il existe des incohérences entre le cadre géographique et les données de la SCP, il peut servir pour la majorité des enregistrements. Cette efficacité peut entraîner les améliorations suivantes :

- Moins de ressources de géocodage manuel seront nécessaires pour que le FCCP et le FCPCEF puissent être disponibles;
- Le délai entre la disponibilité des résultats du Recensement de 2006 et la diffusion des produits fondés sur les codes postaux (FCCP et FCPCEF) peut être réduit.

Il existe des différences de dates de référence entre les données du recensement (disponibles une fois tous les cinq ans), les fichiers de la Société canadienne des postes (disponibles une fois par mois) et les fichiers du réseau routier, qui sont périodiquement mis à jour. Cela peut entraîner des difficultés dans la création de l'association ou du lien dans le cadre du processus de géocodage. Lorsqu'il manque des données à jour dans les fichiers du réseau routier, on résout habituellement la question en communiquant avec la municipalité appropriée et en obtenant l'information. Il peut s'agir d'un processus long, qui peut imposer un fardeau à la municipalité. Une façon de résoudre ce problème consisterait à échelonner le géocodage, c'est-à-dire que l'on coderait en fonction de la SDR ou de l'AD uniquement en l'absence de données dans le Fichier

du réseau routier, et que l'on coderait selon des niveaux plus précis, au fur et à mesure que les renseignements deviennent disponibles. À l'heure actuelle, une fois qu'un code postal est géocodé, nous ne le récupérons jamais pour tenter d'en améliorer la précision. En fait, il se peut qu'un code postal ne soit jamais géocodé à nouveau, à moins que la Société canadienne des postes n'indique un changement de service pour ce code postal. Grâce au géocodage automatisé, nous pourrions envisager des améliorations de la précision avec chaque version du Fichier du réseau routier.

B. Exactitude et précision

L'exactitude du processus de géocodage se mesure par le degré de conformité du géocodage selon la ou les régions géographiques normalisées appropriées. La précision peut être mesurée en fonction de la correspondance entre le nombre de côtés d'îlot et d'îlots de diffusion codés et les tranches d'adresses desservies par le code postal.

Une cote de qualité au niveau de l'enregistrement peut être créée, sur la base du processus de géocodage, afin d'indiquer le niveau de confiance de ce dernier. L'accès à plus d'une source d'information pour le géocodage permet en outre l'élaboration d'indicateurs de qualité. Il existe une façon d'attribuer l'indicateur de qualité, à savoir élaborer un indicateur à trois caractères fondé sur le processus de géocodage proprement dit. Les caractères devraient être facilement compréhensibles, « AAA » laissant supposer que la qualité du géocodage a été confirmée par plus d'une source et a été jugée la meilleure, et « CCC » indiquant que le géocodage constitue une approximation. Le tableau suivant illustre le concept proposé.

Indicateurs de qualité pour les enregistrements de CPTypes 1 et 2

L'indicateur de qualité (IQ) final est élaboré sous forme de concaténation, par exemple :

IQ = IQ_1 ¦ IQ_2 ¦ IQ_3

Un IQ = AAA correspond au meilleur géocodage possible.

Indicateur de qualité (IQ_1)

Cet indicateur concerne la qualité du secteur général en fonction duquel le géocodage a été effectué. Il devrait constituer une indication de la certitude du géocodage selon la subdivision de recensement appropriée (et les parties appropriées de cette SDR).

A : Secteur de recherche fondé sur le secteur de service de la RTA, et la correspondance entre la collectivité de la SCP et la SDR

B : Secteur de recherche fondé sur le secteur de service de la RTA seulement

C : Secteur de recherche fondé sur la correspondance entre la collectivité de la SCP et la SDR seulement

Indicateur de qualité (IQ_2)

Il s'agit d'un indicateur de l'appariement des rues pour les enregistrements de CPTypes 1 et 2.

A : Appariement trouvé selon le nom de rue, le type et la direction uniformisés dans le Fichier du réseau routier

Appariement trouvé selon le nom de rue et le type, pour seulement une rue, celle-ci ne comportant pas de direction dans le Fichier Recherche d'adresses ou le Fichier du réseau routier

B : Appariement d'une seule combinaison de rue et de type à l'intérieur du secteur de recherche

C : Appariement satisfaisant par rapport au nom de rue seulement, ou au nom de rue et à la direction

Indicateur de qualité (IQ_3)

L'IQ a été attribué de la façon suivante dans les appariements de tranches d'adresses et du Fichier du réseau routier.

A : Parité appariée pour les deux adresses du Fichier du réseau routier

B : Parité appariée pour une adresse seulement du Fichier du réseau routier

C : Parité non appariée, mais chevauchement des tranches dans le Fichier du réseau routier

Cet indicateur de qualité s'applique au processus d'appariement des adresses municipales pour les CPTypes 1 et 2 (environ 97 % des enregistrements du Fichier Recherche d'adresses). Un indicateur de la qualité des données pourrait aussi être fourni de façon similaire pour les codes postaux codés directement selon une installation.

Lorsque le géocodage est effectué sur la base des données du recensement seulement, une mesure de stabilité pourrait être créée, à partir de la comparaison de la région géographique déclarée dans le recensement précédent. Cette validation ne constitue pas une mesure directe de l'exactitude, mais fournit une indication du degré de confiance du géocodage.

Indicateurs de qualité pour les enregistrements de CPTypes 3, 4 et 5

L'indicateur de qualité (IQ) final est élaboré sous forme de concaténation, par exemple :

IQ = IQ_1 ¦ IQ_2

Un IQ = AA correspond au meilleur géocodage possible. Il n'y a pas de IQ_3 disponible pour les CPTtype 3, 4 et 5, étant donné qu'il n'existe pas de données sur les tranches d'adresses municipales décrivant le secteur de service, et que le géocodage n'est pas effectué selon la même précision.

Indicateur de qualité (IQ_1)

Cet indicateur concerne la qualité du secteur général en fonction duquel le géocodage a été effectué. Il devrait constituer une indication de la certitude du géocodage selon la subdivision de recensement appropriée (et les parties appropriées de cette SDR). IQ_1 est élaboré exactement de la même façon que pour les CPTypes 1 et 2.

A : Secteur de recherche fondé sur le secteur de service de la RTA et la correspondance entre la collectivité de la SCP et la SDR
B : Secteur de recherche fondé sur le secteur de service de la RTA seulement
C : Secteur de recherche fondé sur la correspondance entre la collectivité de la SCP et la SDR seulement

Indicateur de qualité (IQ_2)

Il s'agit d'un indicateur de la stabilité de la définition des limites de ces secteurs de service. IQ_2 est élaboré différemment que pour les CPTypes 1 et 2, aucune donnée d'adresse municipale n'étant disponible pour le géocodage.

CPTypes 3, 4, 5 (emplacement du secteur de service du code postal)

A : Secteur stable (c.-à-d. géocodé selon la même AD) sur la base des données du recensement courant et du recensement précédent
B : Secteur fondé sur les données du recensement courant et vérifié par téléphone auprès de la municipalité
C : Secteur fondé sur les données du recensement courant seulement

CPType 3 (emplacement de l'installation)

A : Adresse du Fichier sur les chefs de ménage géocodée selon le Fichier du réseau routier et appariement avec l'AD du fichier des emplacements de la SCP[23]
B : Adresse du Fichier sur les chefs de ménage géocodée selon le Fichier du réseau routier
C : Installation localisée selon le fichier des emplacements de la SCP seulement

La précision du secteur codé peut être mesurée dans le contexte du niveau géographique servant au codage, ainsi que du nombre de régions géographiques normalisées codées. Cela est particulièrement pertinent pour le géocodage fondé sur les adresses municipales. Par exemple, les secteurs de service des enregistrements de CPTypes 1 et 2 sont définis selon les tranches d'adresses. Cela signifie qu'au niveau le plus précis, le géocodage sera effectué selon le côté d'îlot, suivi par l'îlot de diffusion et par l'AD, cette dernière constituant le niveau le moins précis de géocodage. Le nombre d'îlots de diffusion servant au géocodage en proportion du nombre de tranches d'adresses pour un code postal permettra aussi d'indiquer s'il y a eu affectation d'îlots de diffusion inutiles à un code postal. Des mesures de la précision peuvent être utilisées dans le

23. Un fichier d'emplacement de l'installation de livraison a été obtenu de la SCP en janvier 2006.

cadre de la vérification du processus de géocodage, ainsi que pour fournir des indicateurs généraux aux utilisateurs des données.

C. Intégralité

Étant donné que les produits fondés sur les codes postaux (FCCP et FCPCEF) fournissent des liens entre les codes postaux et les régions géographiques normalisées, l'intégralité de la couverture de ces produits peut être mesurée de la façon suivante :

1. Le degré de codage de tous les codes postaux et de leur type selon les régions géographiques normalisées.
2. La mesure dans laquelle chaque région géographique normalisée habitée se voit affecter un code postal.

Le changement de méthodologie proposé ici permet de coder les codes postaux de CPTypes 3 et 5 en fonction des AD habitées, lorsqu'aucun itinéraire de facteur ou service d'itinéraire n'est disponible. Cela donnera lieu à une couverture plus complète du service postal offert au Canada.

D. Pertinence

Les ventes ainsi que l'utilisation des FCCP et des FCPCÉF à Statistique Canada montrent que ce produit est utile. Toutefois, un indicateur du fichier, l'indicateur de lieu unique (ILU), sert à la répartition de la population ou des logements, même s'il n'est pas véritablement destiné à cette fin. L'ILU est conçu pour fournir un point représentatif pour un code postal, aux fins de la mise en correspondance des points d'identification du code postal. Toutefois, cet indicateur est aussi utilisé pour l'affectation des données. Il se peut qu'une région géographique normalisée (îlot de diffusion ou AD) ne soit pas représentative de toute la population ayant accès à ce code postal. Par exemple, un secteur de service de code postal peut englober trois îlots de diffusion, et il se peut qu'un îlot de diffusion ne soit pas plus représentatif que les autres. Il est recommandé que la répartition des données se fasse au moyen du logiciel statistique offert par Statistique Canada (FCCP+) ou d'un autre processus de répartition fondé sur un fichier pondéré de la population des AD[24].

E. Intelligibilité et métadonnées

Des métadonnées peuvent être fournies au niveau de l'enregistrement, afin de permettre aux clients de mieux comprendre le couplage des enregistrements. Ces métadonnées contribueront aussi à l'évaluation interne des améliorations qui pourraient être nécessaires dans le géocodage. L'indicateur de qualité (IQ) a été abordé précédemment (dans la section 3.3). Les autres métadonnées proposées au niveau de l'enregistrement sont décrites ci-après.

- Méthode de codage

La méthode de codage peut être pertinente selon la qualité du couplage des enregistrements. On propose ce qui suit pour cet indicateur :

1. Codage manuel confirmé au moyen de cartes
2. Codage manuel confirmé par un appel à la municipalité ou aux autorités provinciales compétentes
3. Géocodage automatisé
4. Géocodage effectué selon les régions géographiques de 2001 et modifié en fonction des régions géographiques de 2006.

24. Le logiciel *Fichier de conversion des codes postaux Plus* (FCCP+) est un produit complémentaire du Fichier de conversion des codes postaux (FCCP). Il s'agit d'un système automatisé qui utilise les codes postaux pour attribuer les régions géographiques du recensement. FCCP+ est fondé sur le Fichier de conversion des codes postaux le plus récent et sur le Fichier de la pondération de la population par codes postaux produit par la Division de la géographie de Statistique Canada. FCCP+ utilise des poids pour affecter les codes postaux liés à des *aires de diffusion* multiples, selon la répartition de la population en fonction d'un code postal donné.

À l'heure actuelle, il est proposé de produire un IQ uniquement pour les enregistrements géocodés de façon automatique. Cet indicateur supplémentaire documente la source des autres types de couplages d'enregistrements. Grâce au contrôle de la qualité de ces autres types de codages, l'IQ pourrait inclure toutes les sources de codage à l'avenir.

- Secteurs de service codés

Le code postal peut être codé en fonction du secteur de service du code postal ou d'une installation de livraison, et ces deux types de codage doivent être distingués. On propose ce qui suit pour cet indicateur :

1. Codé selon le secteur de service du code postal
2. Codé selon l'installation de livraison

- TML avec modifications

À l'heure actuelle, lorsque le TML est absent, Statistique Canada attribue un « W ». Ce « W » comprend à la fois les régions rurales et les petites régions urbaines. L'utilisateur peut être en mesure de mieux utiliser le TML si une distinction est faite entre elles. Une autre lettre, comme « S », pourrait être utilisée pour faire une distinction pour les régions de livraison rurale, lorsqu'un enregistrement est de CPType 2 et que la livraison à une adresse municipale est disponible.

Les règles de géocodage et des sources bien définies pour le géocodage permettent la création d'indicateurs de qualité et d'autres métadonnées qui peuvent renseigner l'utilisateur, ainsi que permettre le contrôle de la qualité des fichiers.

Conclusion

Le présent document visait à : 1) étudier les concepts et descripteurs de codes postaux utilisés par la Société canadienne des postes; 2) établir un cadre conceptuel définissant le rapport entre les codes postaux et les régions géographiques normalisées; et 3) définir par la suite des règles opérationnelles pour le couplage des codes postaux et des régions géographiques normalisées.

Les codes postaux peuvent être liés à des régions géographiques normalisées grâce aux secteurs de service de code postal. Même si, dans certains cas, les données sur les adresses disponibles auprès de la Société canadienne des postes sont suffisantes pour géocoder le code postal en fonction du Fichier du réseau routier, des renseignements supplémentaires sont nécessaires dans nombre de cas pour procéder au géocodage. Le recensement fournit un moyen idéal pour faire une distinction entre le secteur de service de la RTA et celui du code postal. Même si cette méthode de géocodage met l'accent sur l'utilisation des codes postaux par les ménages, des codes postaux d'entreprises peuvent aussi être *géocodés* de cette façon[25].

Pour le géocodage du code postal, la première étape consisterait à déterminer le secteur de service de la RTA, sur la base du recensement, et à la compléter par un fichier de concordance de la collectivité de la SCP et de la SDR. Cela permettrait d'établir le secteur de recherche à l'intérieur duquel le code postal peut être géocodé. La différenciation des empreintes géographiques des codes postaux selon le type permet un meilleur codage selon les secteurs de service de code postal. Un processus de géocodage automatisé fondé sur cette méthode permettra d'assurer l'efficacité du géocodage. Un système efficace de géocodage peut être utilisé au besoin pour profiter des mises à jour des codes postaux, ainsi que des améliorations du réseau routier. En dernier ressort, le cadre conceptuel abordé ici fournit les règles pour un processus de géocodage efficace, y compris des métadonnées sur la qualité de l'enregistrement *géocodé*.

25. Les auteurs suggèrent que l'utilisation des codes postaux par les entreprises, ainsi que les comparaisons de l'empreinte du *secteur de service du code postal* pour le Recensement de 2001 et de celle du Recensement de 2006, soient explorées dans un document de travail subséquent.

Annexe A – Structure du code postal

Le code postal est un code alphanumérique à six caractères qui sert à indiquer la destination de chaque pièce de courrier expédiée au Canada. Les caractères alphanumériques sont disposés selon la forme ANA NAN, le « A » désignant un caractère alphabétique et le « N », un caractère numérique (par ex., K1A 0T6). Le code postal a recours à 20 caractères alphabétiques et à 10 caractères numériques. Six caractères alphabétiques (D, F, I, O, Q et U) ne sont pas utilisés à l'heure actuelle.

Le premier caractère du code postal est assigné par ordre alphabétique, de l'est à l'ouest du Canada, et représente une province, un territoire ou un secteur important entièrement inclus dans les limites d'une province.

Les trois premiers caractères du code postal (ANA) correspondent à des régions précises et stables appelées « régions de tri d'acheminement » (RTA). Une RTA désigne une région particulière à l'intérieur d'une région géographique importante ou d'une province/territoire.

Province/Territoire/Région	Premier caractère du code postal
Terre-Neuve-et-Labrador	A
Nouvelle-Écosse	B
Île-du-Prince-Édouard	C
Nouveau-Brunswick	E
Est du Québec	G
Montréal métropolitain	H
Ouest du Québec	J
Est de l'Ontario	K
Centre de l'Ontario	L
Toronto	M
Sud-ouest de l'Ontario	N
Nord de l'Ontario	P
Manitoba	R
Saskatchewan	S
Alberta	T
Colombie-Britannique	V
Territoires du Nord-Ouest et Nunavut	X
Territoire du Yukon	Y

Annexe B – Qualité des données d'information et qualité des données spatiales

Le but premier de la politique de Statistique Canada visant à informer les utilisateurs sur la qualité des données et la méthodologie est de faire en sorte qu'ils aient accès à l'information dont ils ont besoin pour juger de la pertinence des données en regard de l'usage qu'ils veulent en faire. Six dimensions de la qualité ont été définies à l'intérieur de ce concept de « l'adaptation des données à leur utilisation » et figurent dans le tableau ci-après :

Six dimensions de la qualité de l'information

1. La ***pertinence*** de l'information statistique fait référence au degré d'adéquation de l'information aux besoins réels des utilisateurs, c'est-à-dire de la mesure dans laquelle l'information apporte des précisions sur les questions qui revêtent le plus d'importance pour les utilisateurs. L'évaluation de la pertinence doit tenir compte des besoins variables des utilisateurs.
2. L'***exactitude*** de l'information statistique fait référence à l'exactitude avec laquelle l'information décrit le phénomène qu'elle était censée mesurer. L'exactitude se caractérise habituellement en fonction de l'erreur dans les estimations statistiques et elle se décompose généralement entre le biais (erreur systématique) et la variance (erreur aléatoire). Elle peut aussi être décrite en fonction des principales sources d'erreur susceptibles de causer l'inexactitude (p. ex., erreurs de couverture, d'échantillonnage, de non-réponse, de réponse).
3. L'***actualité*** de l'information statistique fait référence au temps qui s'écoule entre, d'une part, le point de référence (ou la fin de la période de référence) auquel se rapporte l'information et, d'autre part, la date à laquelle l'information est diffusée. Il s'agit généralement de parvenir à un équilibre entre l'actualité et l'exactitude des données. L'*actualité* de l'information influe en retour sur sa *pertinence*.
4. L'***accessibilité*** de l'information statistique fait référence à la facilité avec laquelle les utilisateurs peuvent y avoir accès, y compris la facilité avec laquelle il peut être établi que cette information existe. Cette dimension couvre également la pertinence de la forme ou du support sur lesquels l'information est disponible; le coût de l'information peut aussi être un facteur d'*accessibilité* pour certains utilisateurs.
5. L'***intelligibilité*** de l'information statistique fait référence à la disponibilité de l'information complémentaire et des métadonnées nécessaires à l'interprétation et à l'utilisation adéquates de l'information. Cette dimension englobe normalement les concepts sous-jacents, les variables et les classifications utilisés, la méthode de collecte des données, ainsi que les indicateurs de l'exactitude de l'information statistique. La présente Politique vise à assurer l'intelligibilité de l'information diffusée.
6. La ***cohérence*** de l'information statistique reflète dans quelle mesure l'information peut être combinée avec succès à d'autres données statistiques, à l'intérieur d'un vaste cadre analytique et dans le temps. L'utilisation de concepts et de classifications types, ainsi que de populations cibles, favorise la cohérence, tout comme l'utilisation d'une méthodologie commune pour diverses enquêtes. La *cohérence* ne signifie pas nécessairement une cohérence numérique totale.

Les éléments de la qualité des données spatiales ont pour objet de fournir des renseignements sur la mesure dans laquelle les données conviennent à un usage particulier, en décrivant quand et comment les données ont été créées et en indiquant la conformité de l'ensemble de données avec d'autres ensembles, ainsi que l'exactitude des données. Les éléments comprennent un aperçu décrivant l'objectif et les utilisations, ainsi que des éléments particuliers de la qualité concernant l'historique, la précision de localisation, la précision des attributs, la cohérence logique et l'intégralité. Les éléments de la qualité des données spatiales figurent ci-après.

Éléments de la qualité des données spatiales

1. Historique

L'historique des données spatiales comprend une description des fichiers d'origine à partir desquels les données ont été extraites ainsi que des méthodes utilisées, y compris la date des fichiers d'origine et les transformations effectuées en vue de la production de la version définitive des fichiers numériques ou des produits cartographiques.

2. Précision de localisation

Indique la précision absolue et relative de la localisation des traits géographiques. La précision absolue s'entend du degré de correspondance entre les coordonnées figurant dans l'ensemble de données et les vraies valeurs ou celles acceptées comme telles. La précision relative s'entend du degré de correspondance entre la localisation relative des traits et leur vraie localisation relative ou celle acceptée comme telle. Les énoncés relatifs à la précision de localisation font état de la qualité du fichier ou du produit final après toutes les transformations dont il a fait l'objet.

3. Précision des attributs

Par précision des attributs, on entend la précision des données quantitatives et qualitatives reliées à chaque trait (tels que la population pour une région urbaine, le nom de rue, le nom et le code de la subdivision de recensement).

4. Cohérence logique

La cohérence logique décrit la fidélité des relations encodées dans la structure des données spatiales numériques.

5. Intégralité

L'intégralité indique dans quelle mesure les traits géographiques, leurs attributs et leurs relations sont présents ou non dans un ensemble de données. Elle comprend aussi des renseignements sur les critères de sélection, les définitions utilisées et les autres règles cartographiques pertinentes.

Annexe C – Collectivité de la SCP et SDR

La principale approche recommandée dans le présent document en ce qui a trait au géocodage consiste à utiliser les secteurs de service pour créer des secteurs de recherche. Toutefois, de tout temps, la principale approche pour le géocodage a consisté à localiser la SDR ou une partie de la SDR représentée par le nom de collectivité de la SCP, puis à procéder au géocodage selon les données sur les adresses. La correspondance entre la collectivité de la SPC et la SDR doit être fondée sur une meilleure base que sur la similarité des noms seulement. Cette approche de géocodage sur la base de la SDR est maintenant recommandée comme approche secondaire. Elle permettra le géocodage dans les cas où il n'a pas été possible de créer des secteurs de service ou de confirmer le géocodage effectué à partir des secteurs de service de RTA.

L'approche manuelle de géocodage des codes postaux à Statistique Canada consiste à commencer par apparier le nom de collectivité de la SCP du Fichier Recherche d'adresses et les données tirées d'un Fichier du réseau routier. À cette fin, on tente d'apparier le nom de collectivité de la SCP et le nom de la SDR. Lorsque cela n'est pas possible, des cartes supplémentaires sont consultées et des appels sont faits aux autorités municipales, afin qu'elles contribuent à ce processus. Comme il est indiqué dans le présent document de travail (dans la section 1.3.1), les noms de collectivités de la SCP peuvent être appariés directement avec les noms de SDR moins de 40 % du temps, et c'est ce qui fait en partie que le géocodage prend du temps et est inexact. Par ailleurs, même lorsqu'un appariement est trouvé pour le nom, le secteur correspondant à la collectivité de la SCP peut être différent des limites de la SDR du même nom.

L'amélioration de l'appariement entre la collectivité de la SCP et la SDR fournirait des renseignements supplémentaires pour le géocodage. Ces améliorations ont pris les formes suivantes :

1. Appariement du nom de collectivité de la SCP directement avec le nom de la SDR, après uniformisation de la graphie.
2. Appariement du nom de collectivité de la SCP et du nom de la SDR d'un recensement précédent, et repérage de la SDR connexe du recensement courant.
3. Appariement du nom de collectivité de la SCP et d'un nom de localité, et repérage de la SDR correspondante du recensement courant.

Étant donné que la SCP continue souvent d'utiliser les anciens noms de municipalités, cette approche permet l'appariement des noms historiques et des noms actuels de SDR. Cet appariement s'est révélé utile pour le processus de géocodage manuel.

Toutefois, cette approche pourrait entraîner le géocodage possible en fonction de deux rues différentes ayant le même nom dans la SDR. (Généralement, la SCP conserve l'ancien nom de municipalité jusqu'à ce que les questions liées aux noms en double soient résolues.) De ce fait, une approche pour l'appariement des anciennes limites de SDR et des régions géographiques courantes (comme l'aire de diffusion ou l'îlot de diffusion) fournirait peut-être un meilleur secteur de recherche pour le géocodage. Cette approche doit être vérifiée, afin de déterminer si les secteurs de recherche créés de cette façon peuvent être utilisés pour le géocodage automatisé.

Étant donné qu'un appariement de collectivité de la SCP et d'une SDR fondée sur le nom seulement peut être inexact du point de vue du secteur, il n'est pas recommandé comme méthode principale de géocodage. Un appariement du secteur de la collectivité de la SCP et de régions géographiques moins grandes du recensement courant pourrait fournir un secteur de recherche qui pourrait servir au géocodage automatisé. Des vérifications sont nécessaires, afin de déterminer dans quelle mesure cette méthode de géocodage peut compléter le géocodage fondé sur les secteurs de service. Le processus de géocodage manuel est actuellement mené de façon indépendante des travaux relatifs au géocodage automatisé. Les améliorations de

l'appariement des collectivités de la SCP et des SDR se sont déjà révélées efficaces pour le processus de géocodage manuel.

Glossaire

Adresse municipale
Les adresses postales se divisent en deux catégories, soit les adresses municipales ou les adresses d'installations postales. Une adresse municipale comprend le numéro de voirie, le nom de la rue, le nom de la municipalité et le code postal. Le nom de la rue peut aussi être accompagné du type et de la direction. L'adresse municipale peut aussi avoir un numéro de bureau ou d'appartement. (Société canadienne des postes, 2005 a, p.22).

Adresse d'installation postale (ou Adresse de succursale)
Les adresses postales se divisent en deux catégories, soit les adresses municipales et les adresses d'installation postale. Une adresse d'installation postale se compose d'une description du type d'installation de livraison, à savoir la poste restante, un numéro de route rurale, un numéro d'itinéraire motorisé ou un numéro de service suburbain, le nom de la municipalité (nom de la collectivité de la SCP), la province et le code postal. Elle peut aussi inclure le type et la désignation de l'installation postale. (Société canadienne des postes, 2005 a, p.22-23).

Aire de diffusion *(AD)*
Une aire de diffusion (AD) est une petite unité géographique relativement stable formée d'un ou de plusieurs îlots de diffusion avoisinants. Il s'agit de la plus petite région géographique normalisée pour laquelle toutes les données du recensement sont diffusées. Les AD couvrent tout le territoire du Canada.

Case postale
Série de cases habituellement encastrées dans le mur d'un bureau de poste. Le courrier y est trié par numéro de case à partir de l'aire de travail du personnel. Pour avoir accès à son courrier, le client utilise la clé qui lui a été remise au moment de la location de la case (Société canadienne des postes, 2007, Glossaire).

Collectivités de la SCP
Le terme « municipalité » dans le Fichier Recherche d'adresses ne correspond pas nécessairement à la municipalité définie par la loi au moment où la Société canadienne des postes publie les données. Il s'agit de noms utilisés au niveau local, afin de prévenir la confusion en ce qui a trait au secteur desservi. Afin d'éviter toute confusion, on utilise dans le présent document le terme « collectivités de la SCP ».

Côté d'îlot
Le *côté d'îlot* correspond à un côté de rue situé entre deux traits consécutifs qui coupent cette rue. Ces traits peuvent être d'autres rues ou des limites de régions géographiques normalisées. Les côtés d'îlot servent à produire des points représentatifs de *côté d'îlot*, qui sont utilisés pour le géocodage et l'extraction de données du recensement lorsque les adresses de voirie sont connues.

Genre de subdivision de recensement
Les subdivisions de recensement (SDR) sont classées en 55 genres, selon les appellations officielles adoptées par les autorités provinciales/territoriales ou fédérales. Il y a toutefois deux exceptions, soit « Subdivision of unorganized (SNO) » à Terre-Neuve-et-Labrador et « Subdivision of county municipality (SC) » en Nouvelle-Écosse, qui sont des régions géographiques équivalant aux municipalités et ayant été créées par Statistique Canada de concert avec ces provinces, aux fins de la diffusion des données statistiques.

Afin de mieux distinguer les SDR les unes des autres, le nom de chaque subdivision de recensement est accompagné d'une indication du genre de subdivision de recensement, par exemple, Granby, V (pour la « ville » de Granby) et Granby, CT (pour la « municipalité de canton » de Granby).

Géocodage
Le géocodage est le processus utilisé pour attribuer un identificateur géographique (code) aux détails cartographiques et aux enregistrements de données. Les géocodes ainsi créés permettent d'apparier géographiquement les données.

Îlot de diffusion *(ID)*
Un îlot de diffusion (ID) est un territoire dont tous les côtés sont délimités par des rues et/ou les limites des régions géographiques normalisées. L'îlot de diffusion est la plus petite unité géographique pour laquelle les chiffres de population et des logements sont diffusés. Les îlots de diffusion couvrent tout le territoire du Canada.

Indicateur de lien unique (ILU)
L'indicateur de lien unique peut être utilisé pour établir un lien biunivoque entre les codes postaux et les aires de diffusion, les îlots de diffusion et les côtés d'îlot. Ce lien peut s'avérer utile pour la cartographie. Cependant, plusieurs aires de diffusion, îlots de diffusion et côtés d'îlots n'étant pas représentés dans l'ILU, il n'est pas recommandé de l'utiliser pour la répartition de la population dans les régions géographiques.

Installation de livraison
L'installation à partir de laquelle provient le courrier et qui comprend entre autres les bureaux de poste, les succursales postales urbaines, les centres postaux communautaires et les postes de facteur (Société canadienne des postes, 2007, Glossaire).

Itinéraire motorisé
Itinéraire de distribution motorisé incluant la livraison du courrier dans les parcs ou les quartiers industriels (Société canadienne des postes, 2007, Glossaire).

Livraison rurale
Le terme « livraison rurale » est utilisé dans le présent document pour désigner les régions desservies par des codes postaux correspondant à des régions rurales, selon la SCP. La livraison rurale est également une étape servant à la préparation du courrier. La classification de la SCP est fondée sur le deuxième caractère du code postal, de la façon suivante :

- code postal urbain : chiffres de 1 à 9 (E2J). Les codes postaux urbains sont généralement desservis par un facteur ou par des boîtes postales communautaires; ou
- code postal rural : chiffre de 0 (zéro) (A0A). Les codes postaux ruraux sont desservis par des chauffeurs-livreurs des routes rurales et/ou des comptoirs postaux.

Toutefois, cela ne signifie pas que les codes postaux ruraux correspondent aux mêmes régions rurales que celles définies selon Statistique Canada.

Livraison urbaine
Le terme « livraison urbaine » est utilisé dans le présent document pour désigner les régions desservies par des codes postaux correspondant à des régions urbaines, selon la SCP. La livraison urbaine est également une étape servant à la préparation du courrier La classification est fondée sur le deuxième caractère du code postal, de la façon suivante : classification de la SCP est fondée sur le deuxième caractère du code postal, de la façon suivante :

- code postal urbain : chiffres de 1 à 9 (E2J). Les codes postaux urbains sont généralement desservis par un facteur ou par des boîtes postales communautaires; ou
- code postal rural : chiffre de 0 (zéro) (A0A). Les codes postaux ruraux sont desservis par des chauffeurs-livreurs des routes rurales et/ou des comptoirs postaux.

Toutefois, cela ne signifie pas que les codes postaux urbains correspondent aux mêmes régions urbaines que celles définies selon Statistique Canada.

Logement
De façon générale, un logement est défini comme un ensemble de pièces d'habitation. On reconnaît deux genres de logements, soit les logements collectifs et les logements privés. Les premiers sont des logements de nature institutionnelle, communautaire ou commerciale. Les autres désignent un ensemble séparé de pièces d'habitation possédant une entrée privée soit à partir de l'extérieur de l'immeuble, soit à partir d'un hall, d'un foyer, d'un vestibule ou d'un escalier commun situé à l'intérieur de l'immeuble. Il faut qu'on puisse emprunter l'entrée menant au logement sans passer par les pièces d'habitation d'une autre personne ou d'un autre groupe de personnes.

Ménage
Un ménage est généralement défini comme une personne ou un groupe de personnes qui occupent le même logement et qui n'ont pas de domicile habituel ailleurs au Canada ou à l'étranger. Il peut s'agir d'un logement collectif ou d'un logement privé.

Poste restante
Service de livraison du courrier offert par les comptoirs postaux aux personnes qui ne louent pas une case postale, sont en voyage, sont sans adresse fixe dans le secteur de livraison par facteurs ou ne peuvent recevoir le courrier selon les modes de livraison habituels. (Société canadienne des postes, 2007, Glossaire).

Région de tri d'acheminement (RTA)
Les trois premiers caractères du code postal (combinaison alphanumérique de « ANA ») correspondent à la région de tri d'acheminement (RTA). Cette dernière est définie de la façon suivante sur le site Internet de la SCP :

> La région de tri d'acheminement ou RTA représente une zone précise au sein d'une province ou d'une zone géographique importante. La région de tri d'acheminement constitue le fondement du tri primaire du courrier d'acheminement (Société canadienne des postes, 2007, section B, Chapitre 3, page 8).

Régions géographiques normalisées
Le terme « régions géographiques normalisées » est utilisé dans le présent document pour désigner les régions géographiques utilisées pour publier les résultats du recensement et incluent les îlots de diffusion, les *aires de diffusion* et les subdivisions de recensement.

Route rurale (RR)
Livraison effectuée par un entrepreneur aux clients demeurant le long ou près de routes déterminées dans des régions bien définies (Société canadienne des postes, 2007, Glossaire).

Secteur de recherche
Le terme « secteur de recherche » est utilisé dans le présent document pour désigner la région géographique à l'intérieur de laquelle le géocodage peut être effectué à partir des noms de rue et des adresses municipales.

Secteur de service de code postal
Un secteur de service de code postal peut être considéré comme correspondant à un secteur où des personnes ont accès au courrier grâce à ce code postal.

Secteur de service de la RTA
Le terme « secteur de service de la RTA » est utilisé dans le présent document pour désigner le secteur où les personnes ont accès au courrier pour les codes postaux dans une même région de tri d'acheminement (RTA).

Subdivision de recensement
« Subdivision de recensement (SDR) » est un terme générique qui désigne les municipalités (telles que définies par les lois provinciales/territoriales) ou les territoires considérés comme étant

des équivalents municipaux à des fins statistiques (p. ex., les réserves indiennes, les établissements indiens et les territoires non organisés).

Unité statistique
Une unité statistique est l'unité d'observation ou de mesure pour laquelle des données sont recueillies ou calculées. L'unité statistique de base pour les statistiques sociales est la personne, qui sert d'unite d'analyse dans le cadre de la plupart des programmes de statistiques sociales. Parmi les unités utilisées dans les statistiques sociales figurent le logement et le ménage.

Références

Société canadienne des postes, (2005 a). Fichiers techniques des données sur les adresses et les codes postaux, janvier 2005.
http://www.postescanada.ca/business/offerings/address_management/can/other_data-f.asp
(accessible le 15 janvier 2007)

Société canadienne des postes, (2005 b). Fichiers techniques des données sur les chefs de ménage, janvier 2005
http://www.postescanada.ca/business/offerings/address_management/can/other_data-f.asp
(accessible le 15 janvier 2007)

Société canadienne des postes, (2005 c). Fichiers techniques des données sur les modes de livraison, janvier 2005
http://www.postescanada.ca/business/offerings/address_management/can/other_data-f.asp
(accessible le 15 janvier 2007)

Société canadienne des postes, 2007. Guide des postes du Canada
http://www.postescanada.ca/common/tools/pg/downloads/downPG-f.aspe.asp
(accessible le 15 janvier 2007)

Statistique Canada, 2003. Glossaire illustré
http://geodepot2.statcan.ca/Diss/Reference/COGG/Index_f.cfm
(accessible le 15 janvier 2007)

Statistique Canada, Fichier de conversion des codes postaux (FCCP), guide de référence, mars 2006.
http://www.statcan.ca/cgi-bin/downpub/listpub_f.cgi?catno=92F0153GIF2006002
(accessible le 15 janvier 2007)

Statistique Canada, Politique visant à informer les utilisateurs sur la qualité des données et la méthodologie (approuvée le 31 mars 2000)
http://www.statcan.ca/francais/about/policy/infousers_f.htm
(accessible le 15 janvier 2007)